

FEATURES	Source	Location/Qualifiers
CSNS016Z/c		
LOCUS		1204 bp DNA linear GSS 26-JUL-1999
DEFINITION		Drosophila melanogaster genome survey sequence T7 end of BAC
ACCESSION		BACN15A12 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION		AL106628
KEYWORDS		AL106628.1 GI:5622852
ORGANISM		GSS.
SOURCE		Drosophila melanogaster.
REFERENCE		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Dipera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
AUTHORS		1 (bases 1 to 1204)
TITLE		Genoscope.
JOURNAL		Direct Submission
		Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
		- Web : www.genoscope.cns.fr)
		Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBRc11.
FEATURES		
Source		1. .1204
		/organism="Drosophila melanogaster"
		/db_xref="taxon:7227"
		/clone="BACN15A12"
		/clone_lib="DrosBAC"
		/plasmid="pBelOBRc11"
		/note="end : 17"
BASE COUNT		298 a 172 c 106 g 316 t 312 others
ORIGIN		
Query Match		7.7% Score 66.6 DB 17; Length 1204;
Best Local Similarity		21.3% Pred. No. 0.025;
Matches 107; Conservative 181; Mismatches 215; Indels 0; Gaps 0;		
Y	22	ATTGTGATGATTTTCGACAAATTAAGACATTTTAAAGATTTGAATTTTAT 81
Db	1200	ADAKMAAAT 1141
Y	82	GTCATATCAAGAGACCTATTTGGAGACGCAATTTCAAGTATCTTGAATAAGTAA 141
Db	1140	AAKKAAT 1081
Y	142	ATTATAATAGTTTAAATACCTTTTAAATTTTCAATTAATAGTATATATACCACT 201
Db	1080	AAKMAAT 1021
Y	202	TTAATTAAGAGCTTTTGAATTAATTTTAAATTTTAAATTTTGAATTTTGAATTT 261
Db	1020	KAKKKDKKDDATDADMAAKAAKKDKRAKKKKKKKKKKKKKKKKKKKKKKKKKK 961
Y	262	TTTCTAGCTTTGGCTCAAGCTGATGTTCTTAAAGCGCTTTTAATTTGGAGCGGAGAAA 321
Db	960	KKDKKAKKKKKKAAADAAAAAAKKKKKKKKKKKKKKKKKKKKKKKKKKKKDDAWA 901
Y	322	AACCTTTTGGCTTATCAATAGTAAAGATCCATTTGGCATTTTATTTGAACCTTT 381
Db	900	ADADKAAAKKAAAKKAAKKKKKADADKAAAKKAAAKKAAAKKAAAKKAAAKK 841
Y	382	TTTTAGGCTTTGAATAGTTCTTTTGGCTCAAGAGATATCTTGGGGGTTTCTTATTC 441
Db	840	AARAAAKKGGKGGKGGAGGAGRAAAAGATTTTAATTTTAAATTTTAAATTTTAA 781
Y	442	TTTGATTTGATGCACTGGTATAGGGTTATATCTTACAGGACCTTATTTAGATATCAAG 501

Db	780	TKAGKMAKKTTRAAKAKDARGDAATTAKAARAGAAATKGGGAARAGGAAGKADGARGAK	721
QY	502	CTCTGATAGAAATGCTCCAAA	524
Db	720	GKAAAKAGATKGAATAATACMWA	698
RESULT 14			
CNS00YXQ			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 288:			
QY	57	AAAGAATTGAAATGCTATTTTATGCAATCAAGAAGCTCATTTGGGAAGCAATTT	116
Db	28	VWMAWMTWMSAHCOTTTTTTTTTRKKTWNTTTTTTTTTTNTNNNTNNNTT	87
QY	117	CAAGTATATCTTGGAAAAAGTTAAATTAAATAGTTTAAAAACCTTTTAAATTCA	176
Db	88	NTNNTCNTNTTNRCCNNNNNTNNNNAAGWACATAAAAAATTTTTTNTNTN	147
QY	177	TTAATAGTATCTAATATACACAGTTTAAATAAGAGTTTATGAAATTAATTTAAT	236
Db	148	TTTTTNNNTNTNTAANNAANNNNTNTTTTTTTTTTNTTTTTTTTTTTTTTTT	207
QY	237	GTTGTTTTGACGCTTTTGTTTTTTCTAGCTTGCTCAAGCTGATGCTAAAAGC	296
Db	208	TTT	267
QY	297	GCTTTAAATTTGGAGCGGAGAAAACTTTTACGCTATGAAACTAGTAAGAAAGATCCT	356
Db	268	TTTTTTTTTTTTTTTTT-----TTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	317
QY	357	ATTGCGCAATTTTATTAACCTTTTTTTTAGCGTTTGAATAGGTTCTTTGCTCAAGA	416

1. 1101

servative 4b; Mls

1
2
3
4
5
6
7
8
9
10
11
12

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

CTAGCCTTTGTGTTTTC

TTGGGAGCCGGAGAAACTT

ATTTTTATTCGAACCTTTT

[illegible]

anomic survey sequence

Drosophila melanogaster.

idae; Drosophilidae;
as 1 to 1101)

91006 EVRY cedex - FH

is constructing a

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
/ad_xiel=caxon: /zz/  
/clone="BACR29B23"
```

205; Conservative 133; Mismatches 319; Indels 6; G

.....

.....

AATTAAGAGCGTTTATGAATAATTGTTCTGCTGGCCCTTGCTT

TTCTAGCTTTGCTCAAGCTGATGATTTCTAAAGCGCTTTAATTTGGGAGCGGAGAA

ACTTTTAGCTTATGAACCTAGTAGAAGATCCTATTGTGCCATTTTATTGAACCTTT

TTTAGGGTTTGGAAATAGGTCCTTTTGGCTCAAGGAGATATCTGGGGTTTCTTATTC

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 104

: | | | : | | | . . . | . | | |
 A A A T A T T T T T W A W T A W A W W T W A A A A T W W W
 A A A T A T T T T T W A W T A W A W W T W A A A A T W W W

WTTTWTWATTAATTTWTTWWAWATTWKKKKKKKAWADTGAGARATKTDKKKKKKK

GS 440

[illegible]

RESULT 10	CNS000SX	LOCUS	DEFINITION	1101 bp	DNA	linear	GSS 03-JUN-1999
CNS000SX	Drosophila melanogaster genome survey sequence TE13 end of BAC # BR02L06 of Rpi-36 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	AL050813	AL050813.1	GI:4930826	GSS.	Drosophila melanogaster.	Drosophila melanogaster.
VERSION	KEYWORDS	SOURCE	ORGANISM				
REFERENCE	AUTHORS	TITLE					
JOURNAL							

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Kazuhiro Oseegawa and Aaron Mammossier in Pletier de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPK1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES	source	Location/Qualifiers
1. .1101		
organism="Drosophila		melanogaster"
/dp.xref="taxon:7227"		
/clone="BACR02106"		
/clone_lib="RPCT198"		
/note="end : TE13"		
26 c	85 g	768 t
77 a		145 others

Query Match 7.8%; Score 67.6; DB 17; Length 1101;
 Best Local Similarity 40.2%; Pred. NO. 0.017;
 Matches 287; Conservative 40; Mismatches 389; Indels 2; Gaps 2.

QY	138	TTAAATTTAAATAGCTTTTAAACCTTTTAAATTCATATATATGCTACTATATAC	197
Db	107	TTTTTTTTCATTTTTTTNNNNNNNNNTTTTTTTTATTTTTTTTTTTTTTTTTT	166
QY	198	AGTTTAAATAAAGAGTTTTTATGATAAATTTTAAATGTGTTTGGTACCTTTTGT	257
Db	167	TT	226
QY	258	GTTTTTTCAGCTTTGCTCAGCTGATGATCTAAAGCGCTTTTAAATTTGGAGCGGA	317

[illegible][illegible]

Qy	798	TGATTGGCATTGTATTTGAATAATTAGAGCTTTCCTTTATATATATATATTTTAT	851
		: : : :	
Dd	766	TTTTTGCT-WTTATRATTTRTAGTGCATTTTAGTKATTATKAGTTTTTTATRW	818
RESULT 11			
CNS003B0/c		1101 bp	DNA linear GSS 03-JUN-1999
LOCUS	CNS003B0		
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR0BE08 of RFLP-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		

ACCESSION	AL064078
VERSION	AL064078.1
KEYWORDS	GI:4941834
SOURCE	GSS.
ORGANISM	Drosophila melanogaster. Drosophila melanogaster. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL	Bp 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Oseguwa and Aaron Mammocci in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw^{sp}, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or

Db	1084	TTTAAATTAATTTTTTTTTTAATTATTTTTTTTATTTTTTTTTTTTTTTTTTANTTTTT	1138
RESULT 8	CNS015WU/c	1203 bp DNA linear GSS 26-JUL-1999	
LOCUS	CNS015WU	Drosophila melanogaster genome survey sequence sp6 end of BAC	
DEFINITION	BACN15E10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AJ106008		
VERSION	ALI06008.1 GI:5619558		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 1203)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	determination of this BAC-end sequence was carried out as part of a collaboration with The European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelosBAC11.		
FEATURES	Location/Qualifiers		
source	1..1203 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACN15E10" /library="DrosBAC" /plasmid="pBelosBAC11" /note="end : Sp6"		
BASE COUNT	288 a 103 c 235 g 253 t 324 others		
ORIGIN			
Query Match	7.9% Score 67.8; DB 17; Length 1203;		
Best Local Similarity	40.2%; Pred. No. 0.016; Mismatches 95; Indels 0; Gaps 0;		
Matches	96; Conservative 48; Mismatches 95; Indels 0; Gaps 0;		
21	AATTTGTAGAAATTTTCGACAATAAAGAATTTTAAAGAATTGAATTGGCTAATTTTA 80		
1132	WATWTATATATWTTATWTTTATATWTTATWTTATWTTTAAAAAMAAAAMATYTTTWAMWA 1073		
81	TGGCAATCAGAAGAGCTCTATTTGGGAGCGGAATTTCCAGTAACTTTGAAAAAAGTTA 140		
1072	TWATMAATWMAAAAWMTTATTAWMAWMATAAAMAWMAWMTWTTATATTAATAAANAATP 1013		
141	AATTTAATAGTTATTTAAAAAACCTTTTAAAAATTCATTAAATATGTTACTATATACCAGT 200		
1012	AATWTAAMAATWTTWTAWMAAAATTTTATTAWMAAAAMAAWMAATATATWTTATATMAAAT 953		
201	TTTAAATAAGAGCTTTTATGATGATAAATTTTAAATGTGTTGTTGCTAGCCCTTTTGCT 259		
952	SATWTATWMAATTTTWTATTAATTTTATTAATTTTWCCTTATATWTCAMCACCOCYYTTMYTWT 894		
RESULT 9	CNS005F3/c	960 bp DNA linear GSS 03-JUN-1999	
LOCUS	CNS005F3	Drosophila melanogaster genome survey sequence TEK3 end of BAC #	
DEFINITION	BACR11P07 of RPEI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AI059925		
VERSION	AI059925.1 GI:4943047		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		

ORGANISM	Drosophila melanogaster
Eukaryota:	Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera:	Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea:	Drosophilidae; Drosophila.
REFERENCE	1 (bases 1 to 960)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aaron Mammosses in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and Est libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm. Location/Qualifiers
FEATURES	source 1..960 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACR11P07" /clone_1lb="RPc1-98" /note="end : TET3"
BASE COUNT	570 a 112 c 37 g 71 t 170 others
ORIGIN	
Query Match	7.8% Score 67.6; DB 17; Length 960;
Best Local Similarity	35.4%; Pred. No. 0.018; Mismatches 369; Indels 5; Gaps 1;
Matches 253; Conservative	88; Mismatches 369; Indels 5; Gaps 1;
Oy	143 TTAAATAGTTTAAAAAACCCTTTTAAATTTCATTATATGTACTATAATACCACTTT 202
Dd	853 TTGTAARKKTKTDTPRCANTTKTKTRGRATATTTTKTGWMTGTATTTKTGTAAGA 794
Oy	203 TAATAAGAGCTTTTATGAATAATTTTTATTTGTGGTGGCAGCCCTTTTGGTGT 262
Dd	793 GTTGARAGTKARTATTAAGATTGCGTGAAGAGCTGKTKRRAAATTTGDTTGGKTTG 734
Oy	263 TTCACGCTTTCGCAAGCGATGATCTTCAAAGCCCTTTTAAATTTGGAGCGGAGAAA 322
Dd	733 DWTWWTTTT-----TTGTTGGTGGARRMWAATKTKSTRAADPLATATAKTGRMAT 679
Oy	323 ACTTTAGCTTATGAACACTAGTAAAGAAGCTCATTTGTGCCATTTTATTAACCTTTT 382
Dd	678 TTTTATTTATTATTTGTGTGTTTAAATTAATGTTTDDTTTATTTTATTTTCKTKT 619
Oy	383 TTTAGGTTTGAATAGGCTTTTTCGCACAGAGATATCTTTGGGGGTTTCTTATCT 442
Dd	618 KAGAGAGIDGGARRKRKTKTTTKTTGGGGAATYTKGTGTTBRATWGARDPTTKTIST 559
Oy	443 TGGATTGTCAGTGTGATATGAGGCTTATATCTTACAGAGCTTATTTAGATATCAAGC 502
Dd	558 TGTTTTTKTKTGTTTTTTGWTTTKTGITRKGTGAARAKRTGTGTGRMTGTGTGRAD 499
Oy	503 TCTTATTAAGAATGCTCCAAAAAGCCGCTTTTAAAGTGACCTTGGGGTAAGGAATGAT 562
Dd	498 AAAAATTTKGADGKGCTGTTGTTTCTTKTKTGTMGAAGKRGKTRKTKGRDTCW 439
Oy	563 GGCAGGTGAGTTCATAGCGCTGTCAACAAGATGTGACAGAAATTAATTCCTGTACAT 622
Dd	438 TTTTGTDTGKGAATTTDTTKTKGADMWTTTGTGKKKAAGRITTTTTTTTTTTTTNT 379
Oy	623 TGTCTAATGATTATTAAGAAAATGAAAAATVAGCCTTAATATAGCTTTTGGAGGTTGA 682


```

Db 1034 TWDTWMDMMWMAATKADTAMTWRTAMRADWAGRDGAGKRDADATDADGACRRGG 975
Oy 296 CGCTTTATTTGGAGCGGAGAAAACCTTTAGCTATGAACTAGAAAGATCC 355
Db 974 RKRKRKRKRODDDKKGGKKKAAKATKMDMDKMDKMDKADKRDADDDGA 915
Oy 356 TATGTGCAATTTTATGAACTTTTGAAGGTTGGAATAGCTTTTGTCTCAAG 415
Db 914 GKDGDGKADDDTDTGKDDDKDKMDMDKAKGTGADATMAAATDMMWMAADW 855
Oy 416 AGATTTCTTGGGGTTTCTTATCTTGATTTGATGACAGTGTATAGGTTATCT 475
Db 854 WTMDAAADWADWADWADWAMKMDAMWAGARTADRDWGRGARGKARRRRRAD 795
Oy 476 TACAGCTTATAGATATCAAGCTCTGATAGATGATGATGATGATGATGATG 535
Db 794 KRDADDDDDATTTTATTTTTRDIDDKWKIDTWTMAADTTMDRDDD-DRDRA 736
Oy 536 GTGACCTTGGGTAAGGAATGATGTTGGAGGCTGCTATGCTGAGCAAGATT 595
Db 725 RKRRTWRKRRKRRRTDMDADADTARDRRRRGGDAGKGGKRRRRRRATWD 676
Oy 556 GACAGAAATTTATTTCCGTTTACATTTGCTATATGTAATAGGAAACGAAATAG 655
Db 675 RTDAMWAAWMTTDTDTDDMDKRRRRRRRRRTTARAAMDMWTKAMWAKMDWK 616
Oy 656 CCTTAATATAGCTTTTGGAGGTTGAGCTGATGATGATGATGATGATGATG 715
Db 615 TRADWMDWADTTDADKADBDKAKARARARARARARARARARARARARAR 556
Oy 716 CGCTCTGGTTGGAATCTTTCACAAAAGATTAATTTTATTTTATTTAAT 775
Db 555 AARAAWMAWMAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 496
Oy 776 GAGTATAGCAATTTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
Db 495 AAAAAAATAATTTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 436
Oy 836 ATTATTTATTTTATTTCT 855
Db 435 WTTWMAATTTATTTTWTWT 416

RESULT 5
CNS04P4P/C 1184 bp DNA linear GSS 24-MAY-2000
LOCUS Tetradon nigroviridis genome survey sequence T7 end of clone
DEFINITION 125P20 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL300850
VERSION AL300850.1 GI:8177869
KEYWORDS GSS: genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetradon.
REFERENCE 1 (bases 1 to 1184)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizesen,C., Winker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
HUMAN gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 1184)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizesen,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Unpublished

```

```

REFERENCE 3 (bases 1 to 1184)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetradon.

FEATURES
source
1..1184
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone="125F20"
/clone_id="G"
/note="Genoscope sequence ID : COBG125DC10LP1-end : T7"

BASE COUNT 423 a 218 c 225 g 253 t 65 others

ORIGIN
Query Match 8.1%; Score 69.6; DB 17; Length 1184;
Best Local Similarity 44.4%; Pred. No. 0.0076;
Matches 120; Conservative 31; Mismatches 119; Indels 0; Gaps 0;

Oy 3 TTTTCATTTGATCCAGATTTTGTAGATTTTGCACAAATTAACATTTATAAGAA 62
Db 1081 TTTTNTATATATATATATATATATATATATATATATATATATATATATAT 1022
Oy 63 TTGAATATGCTAATTTATGCAATCAAGAGCTATGGAAGCAATTCACGTA 122
Db 1021 TTTTAAATTTTATTTTAAATTTAAATTTAAATTTTATTTTAAATTTATTT 962
Oy 123 ATACTTAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTAAATTCATTAATA 182
Db 961 ATWTTTATATAATTTAAATTTAAATTTTAAATTTTAAATTTTAAATTTT 902
Oy 183 TGTCTATATATACAGATTTTAAAGAGCTTTTAAATTAATTTTAAATTTGCT 242
Db 901 TATTTTATATTTTAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAAT 842
Oy 243 TTGCTAGCCTTTTGTGTTTCTAGCTTT 272
Db 841 TWTTTTAAATTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 812

RESULT 6
CNS00DKY 928 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL071865
VERSION AL071865.1 GI:4948170
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Acanthopterygii; Drosophila.
REFERENCE 1 (bases 1 to 928)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999)
COMMENT Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org/The BDGP Drosophila
melanogaster BAC library was prepared by Kazuhiro Oseogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the

```

[illegible][illegible]

ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 1101

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="BACR48p19"

/clone_lib="RPCT-98"

/note="end : TET3"

BASE COUNT 469 a 6 c 69 g 151 t 406 others

ORIGIN

Query Match

Best Local Similarity 24.4%; Pred. No. 2.1e-05;
Matches 176; Conservative 216; Mismatches 328; Indels 2; Gaps 1;

```

38 ACAATTAAGACATTTTAAAGAAATGTAATTTTGTCAATCAGAAGC 97
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
98 TCTATTGGAGCGAATTCAGTAATCTTGAAGAACTTAATTAAGTTAA 157
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
442 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
158 AAACCTTTTAAATTCATTAATGTAATTAATCAAGTTTAATAAGGTTT 217
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
502 TTTTWTWTTTDAADKMAAMDTTWTWTTWTTWTTWTTWTTWTTWTT 561
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
218 TATGATTAATTTTATGTTTGTGCTAGCCTTTGCTTGTGCTGCTCA 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
562 WTKAKTAATAAAMWTTTAAATTTTAAATTTTAAATTTTAAATTTT 621
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
278 AGCTGATGATCTAAAGCGCTTAAATTTGAGCGGAGAAACTTTTAC 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
622 TTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTA 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 AACTAGTAAGAATCCATTTGTCATTTTATGAACCTTTTAAAGGTTG 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
682 AAAMWTTTDTWMAAMWTTTAAATTTTAAATTTTAAATTTTAAAT 741
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
398 AGCTGCTTTGCTCAAGAGATATCTGAGGCTTTCTTATCTTGATTC 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
742 KKRKRGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 801
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
458 TGGTATAGGCTTAATCTTACAGAGCTTATTAATCAAGCTCTGATGA 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
802 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 861
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
518 TCAGAAAGCGCTTTAAAGGAGCTGAGGAGTAAGTAATGTCAGAGT 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
862 KKKKKKKG - GKKDAAKKKKKKKKKKKKKKKKKKKKKKKKKKK 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 TATGCTGTGACAGATGACAGAAATTAATTCGTTTACATTTGCTAAT 637
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
920 KKKKGGGGGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 979
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
638 TAGGAACGCAAAATAGCTTAATAGCTTTTGAAGGTTTGAGCTTAT 697
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
980 WTDAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1039
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
698 TAATAGGCGCAAGCTGCTGCTTGGGTTTGAATCTTCAAAAAAGT 757
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1040 KKTGKDKGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGK 1099
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
758 TT 759
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1100 KK 1101

```

RESULT 2
CNS035N7 576 bp DNA linear GSS 15-MAY-2000
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence pUC-ori end of clone
214A06 of library G from Tetraodon nigroviridis, genomic survey
sequence.

FEATURES

source

1. 576

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="214A06"

/clone_lib="G"

/note="Genoscope sequence ID : C0AG214BA03SP1-end ;
pUC-ori"

BASE COUNT 62 a 5 c 18 g 444 t 47 others

ORIGIN

Query Match

Best Local Similarity 44.3%; Pred. No. 0.00032;
Matches 162; Conservative 31; Mismatches 173; Indels 0; Gaps 0;

```

112 AATTCAAGTAATCTTGAAGAAAGTAATTAATAGTTTAAACCTTTTAA 171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 AATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
172 TTTCAATTAATCTTACTAATATACAGTTTAAAGAGCTTTTATGAAT 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
108 TTTTWTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTT 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 TAATGTGCTTTGCTAGCCTTTGCTTTTCTGCTGCTGCTGCTGCT 291
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 TAAWTTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTT 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
292 AAAGCGCTTAAATTTGAGGAGGAGAAACTTTTAAAGTATGAAGT 351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
228 TTTTWTWTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTT 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
352 ATCCATTTGTCATTTTAAAGCTTTTAAAGCTTTTAAAGCTTTT 411
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
288 GGGTTTWTWTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
412 AAGGATATATCTTGGGGGTTTCTTATTTCTTGATTTGATGAGTT 471
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 01:16:47 ; Search time 1209.71 Seconds
(without alignments)
11540.380 Million cell updates/sec

Title: US-09-508-487-20

Perfect score: 862
Sequence: 1 gatttttcattgatccacg.....atatattctctctgctaa 862

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.4	9.8	1101	17	CNS00L72
2	78	9.0	576	17	CNS0035N7
3	77.8	9.0	1101	17	CNS0100X
4	75.6	8.8	1101	17	CNS0039G
5	69.6	8.1	1184	17	CNS04P4P
6	69.2	8.0	928	17	CNS00DKY

7	68.2	7.9	1167	17	CNS07360	AL427102 clone BAO
8	67.8	7.9	1203	17	CNS015W3	AL106008 Drosophila
9	67.6	7.8	960	17	CNS005P3	AL059925 Drosophila
10	67.6	7.8	1101	17	CNS0005X	AL050813 Drosophila
11	66.8	7.7	1101	17	CNS0003B0	AL064078 Drosophila
12	66.6	7.7	1101	17	CNS0003BL	AL069706 Drosophila
13	66.6	7.7	1204	17	CNS016E2	AL106628 Drosophila
14	66.4	7.7	1204	17	CNS000YX	AL096968 Drosophila
15	66	7.7	900	17	CNS03M11	AL250174 Tetradon
16	65.6	7.6	524	17	CNS01090	AL167341 Tetradon
17	65.4	7.6	570	12	BG680408	BG680408 1024070D0
18	65.2	7.6	1101	17	CNS000E07	AL069440 Drosophila
19	64.8	7.5	996	17	CNS000F0H	AL071063 Drosophila
20	64.6	7.5	862	17	CNS029ER	AL187164 Tetradon
21	64.6	7.5	1225	17	CNS0161D	AL106171 Drosophila
22	64.6	7.5	1303	14	BM924877	BM924877 AGENCOURT
23	64.4	7.5	893	17	CNS013XE	AL103436 Drosophila
24	64.4	7.5	915	17	CNS008BP	AL051471 Drosophila
25	64.2	7.4	1094	17	CNS012F2	AL101513 Drosophila
26	64	7.4	1135	17	CNS033GO	AL226115 Tetradon
27	64	7.4	1200	17	CNS016CO	AL106578 Drosophila
28	63.8	7.4	669	17	CNS010M7	AL099145 Drosophila
29	63.6	7.4	916	17	B09084	B09084 F6L16-T7 IG
30	63.2	7.3	865	17	BH180441	BH180441 016_H_20-
31	63.2	7.3	865	17	CNS07MAN	AL617393 T3 end of
32	63.2	7.3	1101	17	CNS00L72	AL078714 Drosophila
33	63.2	7.3	1201	17	CNS0167M	AL106386 Drosophila
34	62.6	7.3	1152	12	BG309087	BG309087 HVSMEC000
35	62.4	7.2	928	17	CNS00DKY	AL071865 Drosophila
36	62.2	7.2	1125	9	AL547503	AL547503 Drosophila
37	62.2	7.2	1201	17	CNS016FX	AL106695 Drosophila
38	62	7.2	936	17	CNS07ANM	AL336808 T7 end of
39	62	7.2	958	17	CNS0074D	AL066801 Drosophila
40	62	7.2	1101	17	CNS00D77	AL075293 Drosophila
41	62	7.2	1101	17	CNS000YWL	AL096927 Drosophila
42	61.8	7.2	661	17	CNS020VJ	AL209800 Tetradon
43	61.8	7.2	836	17	AZ528131	AZ528131 ENTICY19TF
44	61.8	7.2	886	17	BH155984	BH155984 ENTPY20TF
45	61.8	7.2	957	17	BH154830	BH154830 ENTIT170TR

ALIGNMENTS

RESULT 1
CNS00L72 1101 bp DNA linear GSS 14-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL078714.1 GI:5102004
VERSION AL078714
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT
- Web : www.genoscope.cns.fr
- Web : www.fruitfly.org
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>
The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

THIS PAGE BLANK (USPTO)

Fri Mar 28 13:43:34 2003

us-09-508-487-20.rng

Page 13

Job time : 246.586 secs

Db 4032 TTTTGTGTTTGGTAAATGGGATGATGTTATATATATTTTGTAGTTATATGT 4091
 QY 499 AAGCTCTGATAGATAGCTCCCAAAAGCCGCTTTTAACTGACCTGGGTAAGGAATGA 558
 Db 4092 TTTTATTTATATTTTATTTTATGATATTTTATGCTATATTTTACAAATTTATAT 4151
 QY 559 TGTGGCAGGTGACGATGCTATGCTGACAGATGACAGAAATTTATTTCCGTTTA 618
 Db 4152 AGTATTCAGATTTTAAATATTTTGTAGTTAAATATATAGAGTTTGTTTTAACTTT 4211
 QY 619 CATTTGCTAATAGTTATATATAGAAACGAAATAGCCCTTAATATAGCTTTGGAGGT 678
 Db 4212 AATTTATGTTAATATATATTTGGAATATTTAAATTTTAAATGCTTAAATGAT 4271
 QY 679 TTGAGCCATGTTTGAATATATATAGGCGCAAGCTAGCCGCTTGGGTTGACATGCT 738
 Db 4272 TTTTGTGTTTGTAGTTTATGCTGCTGATAGATAGATGTTGTTATTTTAACTA 4331
 QY 739 TCAAAAAAGTTATTTATTTTATTTTATTTTAAATGAGTGAAGCAATTTGTATTT 798
 Db 4332 AATGAAAGATTTTAAATTAAGTAATTTAGTTTATTTTGAATATTTAGTTTATTT 4391
 QY 799 GATTCGCTATTTGTAATGAAATAGAGCTTTGTTTATTTATTTATTTT 849
 Db 4392 TATTTATTTATGATTTTGAATTAATTAATTTGATTTTGTGATTTT 4442

RESULT 15
 AAS45410
 ID AAS45410 standard; DNA; 11036 BP.

XX AAS45410;
 AC 18-DEC-2001 (first entry)
 DT
 XX

Chemically pretreated genomic DNA associated with cell cycle #58.

XX Cell cycle: human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Levy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.

XX Homo sapiens.
 OS
 XX MO200168911-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02945.

XX 15-MAR-2000; 2000DE-1013847.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2001-602751/68.

XX Designing primers and probes for analysing diseases associated with

XX cytosine methylation state e.g. arthritis, cancer, aging, PT

XX arteriosclerosis comprising fragments of chemically modified genes

XX associated with cell cycle -

XX Claim 1; SEQ ID No 115; 28pp; English.

CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing
 CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers.

SO Sequence 11036 BP; 3049 A; 312 C; 2598 G; 5070 T; 7 other;

Query Match 7.6%; Score 65.2; DB 22; Length 11036;
 Best Local Similarity 46.9%; Pred. No. 0.0049;
 Matches 348; Conservative 0; Mismatches 378; Indels 16; Gaps 4;

QY 119 AGTAACTTTGAAAAAAGTTAAATTAATGATTTTAAAAACCTTTTAAATTCAT 178
 Db 6906 AGTGTATATTTTGTGTTTAAATTTTAAATTTGATTTTAAATTTATTTAT 6965
 QY 179 AATATGTTACTAATTAACACGTTTAAATGAAGAGTTTATGATTAATTTAATG 238
 Db 6966 TGTGTTTTTTAGGTTAAATTTATTTGATTTGTTTTTTTGAAGTTTCGATTTAT 7025
 QY 239 TGTGTTGCTAGCTTTTGTGTTTCTGTTTCTGCAAGCTGATTTTAAAGCGC 298
 Db 7026 TAAACGTAATTAATTTAGTGTGTTTATGCTATATATATGATGAGTTTATGAT 7085
 QY 299 TTTTAATTTGGAGCGGAGAAAAAAGTTTATGCTTAAGCAAGTAAGAAAGATCCAT 358
 Db 7086 TTTGAAATATTTTGTGTTTGAATAGTAT---TTATTAAGATTAAGATTAAGATAT 7141
 QY 359 TGTGCAATTTTATTTGAACCTTTTATAGGTTTGAATAGCTTCTTCTCAAGAGA 418
 Db 7142 TTTATATATGTTATGATTTTATTTTATTTTATTTTATTTTATTTATTTATTT 7201
 QY 419 TATCT-----TGGGGGTTTCTTATTTGTTGATTTGATTTGATTTGATTT 468
 Db 7202 GATTTTATGTAATTAATGTTGTTGTTTATTAATTAATTAATTAATTAATTT 7261
 QY 469 TAAATCTTACAGAGCTTATTTAGATTAATCAAGCTTTGATTAAGATCTTCAAAAGCG 528
 Db 7262 ATTTAGTTTATGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7321
 QY 529 CTTTATGAGCTTTGGGTTAAGGA-ATGATTTTGGCAGCTGCACTTACTATGCTGTG 587
 Db 7322 TTTGTTTTTTTATTTTAAATGAGCTTTTGTGTAATTAAGTAAGTAAGTAAGT 7381
 QY 588 ACAAGATTTGACAGAAATTAATTTATTTCCGTTTACATTTGCTATTTAGTTAATGAAACTG 647
 Db 7382 AATAGTTATTTTAAATGTTGTTTATTTATTTATTTATTTATTTATTTATTTATTT 7441
 QY 648 AAAATAGCCTTAATATAGCTTTTGAAGGTTTGAAGCTTATTTATTTATTTATTT 707
 Db 7442 AGAAGTTTGTGTTGAGGTTGCGAAATATTTAGTTAGTATTTGATTTATTTT 7501
 QY 708 CAAGCTAGCGCTTTGGTTGAACATCTTTCAAAAAAGTTTATTTATTTATTTATTT 767
 Db 7502 TTTTATTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7560
 QY 768 ATTTAAATGAGTATGATTTTATTTATTTGATTTGCTATTTATTTATTTATTTATTT 827
 Db 7561 ATTTTAAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 7620
 QY 828 TTTGTTTATTTATTTATTTATTT 849
 Db 7621 TTTGTTATTTATTTATTTATTT 7642

Search completed: March 27, 2003, 01:32:16

QY	51	TTATATAAAGAAATGGAATTCGTAATTTATGTGTCAAATCAAGAGCTGTATGGGAAGC	110
Db	3973	TAAATATATTAATTTTAAAGGAGAAATTCGTTAAATAAAGGGGTATACATTTTATG	4032
QY	111	GAATTCACAGTAATCTTGTGAAAAAGTAATTTAATAGTTTTTAAAAACCTTTTAA	170
Db	4033	TAAATTTGAATTTTAAATGAGGTAGTATTTAAATTTTAAATTTTAAATTTTGTG	4092
QY	171	ATTTCATTAAATAGTACTATATACACGTTTAAATAAGAGTTTATGA-----A	223
Db	4093	ATTTTATGTTTATATTTAGGTTATATGATGTAAAGGTGGGTTTATGGGTTTTTGA	4152
QY	224	TAAATTTTATATGTTTGTGTCACCTTTGTGTTTTTTCAGCTTTGCTCAAGCTGA	283
Db	4153	TATTTTGTGTTTTGTGTTTTGTAGAGTATAGTTTTTTTTTACGTGTTTTTATGGTGA	4212
QY	284	TGATTCATAAAGCGCTTTTAAATTTGGAGCGGAGAAAACTTTAGCTTATGAACATG	343
Db	4213	TATGATGCTTTTGAGTTTTTAAAGTATAGTATGATGTCAGTGTGCGGTACAAATATTT	4272
QY	344	TAGAAAGATTCCTATTTGTCCATTTTATTTGAACCTTTTATAGGTTTGGATAGTTC	403
Db	4273	TAGGAGCAGAGATGCTGTTTTTTTTTTTGTAGTTTTATAGTAGTATTAGGGGGA	4332
QY	404	TTTTGCTCAGAGAGATATCTGTGGGGTTTTCTATCTTGATTTGATGACGTGTGAT	463
Db	4333	TTTTGTGTGGGGTTTTAATTTTATTTTATATGTTTTGTTTTTTTGAATGTGTGTTTT	4392
QY	464	AGGGTATATCTTACACAGACCTTATTTAGATATCAAGCTCTGTATAGATAGTCCAA	523
Db	4393	TTACATTTTTTTTTTTTTTTTTTATATAGATTTTAAATATTTTTTTTTTAAGTTTAAAT	4452
QY	524	AGCCGCTTTTAAAGTGCCTTGGGTTAAGGAAATGATGTTGGCAGGTGCAATATGAGC	583
Db	4453	TTTTATGATTTTTTAGGGAAGGGTAAATATGTTATGTTTTTTGTTAAGTATATAG	4512
QY	584	TGTACAAAGTTACACAGAAATTTATTTCCGTTTACATTTGCTAATAGTTAATAAGAA	643
Db	4513	---GATTTATTTTATTTTAACTTTATATAGATTTTAAATTTTATTTTGGATTTATATAG	4562
QY	644	ACTGAATAATAGCCTTAAATATACCTTTTGGAGGTTTGAGCCAGTTTGTATTAATAT	703
Db	4570	TGTGATTTTATATGTTATATTTATTTATATAGATTTTGGTTAAATATATTAAAGAGTTT	4622
QY	704	GGGCCAAGCTAGCGCTCTGTGGTTGACATCTTCAAAAAAGTATTAATTTTATTT	763
Db	4630	TAGGAATTTTAAATTTTTTTTGATTTTTTATTTTTTGAATTTTTTATATTTTAAATTT	4682
QY	764	TATATATTAANAATAGATATAGCAATTTTGTATTTGATTTGCTCATTTGATAATTA	823
Db	4690	TTGTTGTTTATTTAGTTTAAAGCTTTTATTTAATTTTTTGGGATTTTATATAGTAATAT	4742
QY	824	GAGCTTTTGTATATATTTATTTATTTATTTATTTCTCTGCTAA	862
Db	4750	TTATTTTGTATTAATTTATTTGATTTATAGTTTATTTTAA	4782
RESULT 14			
ABL/70480			
ID	ABL/70480 standard; DNA; 9742 bp.		
XX	ABL/70480:		
XX	01-JUL-2002 (first entry)		
DE	Chemically treated cell signalling DNA sequence complementary to#185.		
XX	Cell signalling: cytosine methylation; cell signalling disease;		
KM	cancer; tumour; cytosstatic; ds.		
XX	Unidentified.		
OS	WO200202807-A2.		

Pd			10-JAN-2002.
Xx	-		
Pf			29-JUN-2001; 2001WO-EP07471.
Xx			
Pf			30-JUN-2000; 2000DE-1032529.
Pr			01-SEP-2000; 2000DE-1043826.
Xx			(EPIG-) EPIGENOMICS AG.
Pa			
Pi	Olek A,	Piepenbrock C,	Berlin K;
Xx			wpi; 2002-154758/20.
Dn			
Rr			
Xx			
Pt	Nucleic acid, useful for diagnosis and therapy of diseases associated with cell signalling e.g. cancer, comprises chemically modified genomic sequences of genes associated with cell signalling -		
Ps	Claim 1; SEQ ID NO 370; 24bp+sequence listing; English.		
Cc	The invention relates to a nucleic acid comprising a sequence of at least 18 bases of a segment of chemically pre-treated DNA of genes associated with cell signalling. The activity of the modified sequences of the invention may be described as cytostatic. The object of the invention is to provide the chemically modified DNA of genes associated with cell signalling, as well as oligonucleotides and/or PNA-oligonucleotides detecting cytosine methylations, as well as a method which is particularly suitable for the diagnosis and/or therapy of genetic and epigenetic parameters of genes associated with cell signalling. The chemically modified DNA provided by the invention is useful for diagnosis and therapy of diseases such as solid tumours and cancer. The sequences given in records AB170111-AB170626 represent chemically pre-treated genomic DNAs of genes associated with cell signalling. Cc Note: The sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.		
Sq	Sequence 9742 BP; 2945 A; 88 C; 1819 G; 4890 T; 0 other; Query Match Best Local Similarity 7.6%; Score 65.8; DB 24; Length 9742; Matches 366; Conservative 0; Mismatches 462; Indels 3; Gaps 2		
Oy	19 ACAATTTGTAGAAATTTCGCACAAATAAAGACAATTAAAGAATGAATGGCTAATTTT Db 3615 ATTAAGTGAAAAAATTTGTATGATATATATAATTTGAATAACGTAAATTAATAATTTTA 	78	
Oy	79 TATGTCCAATCAGAACACTCATTTGGGAGCGAATTCCAGTAATFACCTTAAAAGT Db 3675 TTATATAAAATTTGGAAAT-AAATTTTAGCAATTTTTATTTTAAAGTATATATGA 	138	
Oy	139 TAAATTTAAATAGTTTTTAAAAACCCTTTTAAATTCATTAATATGTCATTAATACGA Db 3734 AAATATTTAATTTAGTTAATAGAAATATTTTTTAAAAATTTGAAAAAGTAA--ATAATAGTA 	198	
Oy	3792 ATTATTTTATATTTGTTTTTTTTTTTTTTATTTTATATTTTGTAAAGAAAAA Db 3792 ATTATTTTATATTTGTTTTTTTTTTTTTTATTTTATATTTTGTAAAGAAAAA 	3791	
Oy	259 TTTTCTTACGCTTTCCTCAACGCTGATTTCTAAAGCGCTTTAATTTGGAGCGGGAG Db 3852 TTATTAATTTTTAGAGGGGTAGAGAGCTAATTTTAAAGATGAAATTTTTCGTTAGTA 	318	
Oy	319 AAAAATCTTTAGCTTATGAACACTACTAAGAAAGATCCATTTGCCATTTTATGAGNC Db 3912 TTCGGTGTAGATGAGTTTAAAGTAAAGCAATTAATATTTAVGAAANAATTAAGATATT 	378	
Oy	3792 TTTTATTTAGGCTTTGCGTTCGTCAGAGAGATATTCGTTGGGGCTTTCTTA Db 3972 ATGCTGTTGAGTAATTTGCTTTTGCAATTTGAAAAAATTTTTTATTTTGTATATTTG 	438	
Oy	439 TTTCTGATTTTGATGCAAGTGTGGTATAGGCTTAATCTTACAGAGACTTATTTAGATATCA 	4031	

xx The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 CC with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or PNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-ABL70626 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

xx Sequence 6012 BP; 1842 A; 43 C; 1153 G; 2973 T; 1 other:

Query Match 7.8%; Score 67; DB 24; Length 6012;

Best Local Similarity 44.4%; Pred. No. 0.0022;

Matches 364; Conservative 0; Mismatches 445; Indels 10; Gaps 2;

51 TTATTAAGAATGCAATTCCTAATTTATGTCACAAATCAAGAAGCTCTATTGGAGAC 110
 Db 3973 TAAATATATTTAATTTAAATGGAGAAATTTGTTAAATTAAGGGCTTATGATTTTATG 4032
 Oy 111 GAATTTCAAGTAATACCTTTGAAAAAGTTAAATTTAAATTTAAATTTAAATTTTAA 170
 Db 4033 TAAGTTGAATTTAATGAGTAGATTTAAATTTAAATTTAAATTTAAATTTTGTG 4092
 Oy 171 ATTTCATTAATATGTTACTATAATACAGTTTATATAAGAGCTTTATGCA-----A 223
 Db 4093 ATTTATGTTTATATTTAGTTATATGTTATATGATGATGAGTGGCTTTTATGGCTTTTGA 4152
 Oy 224 TAAATTTTAAATTTGTTTCTTGTAGCCTTTTGTGTTTCTTGTAGCTTTGTCAGCAAGCTGA 283
 Db 4153 TATTTTGTGTTTGTGTTTGTAGCTTATAGTTTGTGTTTGTGTTTGTGTTTGTGTTG 4212
 Oy 284 TGATTTCTAAAAAGCCTTTTAAATTTGGAGCGGAGAAAACTTTTACGTTATGAACATG 343
 Db 4213 TATTGAGTGTGTTGTTTAAATTTTAAATGATGATGATGATGATGATGATGATGATGAT 4272
 Oy 344 TAGAAGAAAGTCCATTTGTCATTTTAAATTTGCAATTTTAAATTTTAAATTTTAAATTTTAA 403
 Db 4273 TAGGAGACAGAGATGATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 4332
 Oy 404 TTTTGTCTCAAGAGATATTTTGGGGGTTTCTTATCTTGGATTGATGACAGTTGTAT 463
 Db 4333 TTTTGTGTTGGGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 4392
 Oy 464 AGGTTAATCTTACAGAGCTTATTTAGATATCAACGCTGTGATGAATGATGCCAA 523
 Db 4393 TTAGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 4452
 Oy 524 AGCCCTTTAAGTACTTGGGAGTAAAGGATGATGTTGGAGGTGACAGTATATGAC 583
 Db 4453 TTTATAGATTTTAAAGGAGGATTAATGTTATGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 4512
 Oy 584 TGTGCAAGATTTGACAGAAATTTATTTCCGTTTACATTTGCTAATTTAGTTATTAAGAA 643
 Db 4513 ---GATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 4569
 Oy 644 ACTGAAAAATAGCTTATATATGCTTTTGGAGGCTTTGACCTTATTTGATATTAATAT 703
 Db 4570 TGTGGATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 4629
 Oy 704 GGGCCCAAGTACGCTTGTGGGTTGAACATCTTTCAAAAAATTTTAAATTTTAAATTTTAA 763
 Db 4630 TAGGAGATTTTAAATTTTGTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 4689
 Oy 764 TATTAATTAATGAGTATGACAAATTTGTATTTGATTTGCTCATTTGTAATTTGAAATTTA 823

Db 4690 TTGTTTGTATTTAGTTTAAAGTTGTTTAAATTTTGTGGCTATTTTAAATTAATTT 4749
 Oy 824 GAGCTTTGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 862
 Db 4750 TTATTTTGTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 4788

RESULT 13

ID ABL31187 standard; DNA; 6012 BP.

XX ABL31187;

XX 23-APR-2002 (first entry)

DE Signal transduction associated gene modified complementary DNA #15.

XX Human; signal transduction associated gene; cytosine methylation state;

KW Cpg island; signal transduction associated disease; solid tumour; cancer;

XX antitumour; cytostatic; mutant; ds.

OS Homo sapiens.

XX Synthetic.

XX WO200200926-A2.

XX 03-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07472.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIC-) EPICENOMICS AG.

XX Olek A. Piepenbrock C. Berlin K;

XX WPI; 2002-147896/19.

PS Claim 1; SEQ ID No 30; 24pp; English.

CC The present invention relates to chemically modified DNA sequences of

CC signal transduction associated genes. The DNA sequences are chemically

CC modified using a solution of bisulphite, hydrogen sulphite or

CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers

CC for detecting the cytosine methylation state (Cpg islands) of these

CC genes, and a method for the diagnosis and/or therapy of genetic and

CC epigenetic parameters of genes associated with signal transduction.

CC The genomic DNA can be obtained from cells or cellular components which

CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,

CC cerebral spinal fluid, tissue embedded in paraffin such as tissue from

CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,

CC histologic object slides, and all their possible combinations. The

CC sequences of the invention are useful for the diagnosis and therapy of

CC diseases associated with signal transduction e.g. solid tumours and

CC cancer. ABL31158-ABL31545 represent chemically pretreated genomic DNA

CC sequences of different genes associated with signal transduction, or

CC their complementary sequences.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from the

CC European Patent Office.

XX Sequence 6012 BP; 1842 A; 43 C; 1153 G; 2973 T; 1 other:

Query Match 7.8%; Score 67; DB 24; Length 6012;

Best Local Similarity 44.4%; Pred. No. 0.0022;

Matches 364; Conservative 0; Mismatches 445; Indels 10; Gaps 2;

Matches 337; Conservative 0; Mismatches 420; Indels 8; Gaps 1;

PD 20-SEP-2001

XX

XX (HUMA-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMUNE INC.
 XX ChOI GH, Erwin AL, Hanson MS, Lathigra R;
 XX WPI: 1999-189980/16.
 DR P-PSDB: AAY19809.
 XX
 XX New isolated *Borrelia burgdorferi* nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by *Borrelia*, particularly Lyme disease
 XX
 XX Claim 1: Page 73; 275pp; English.
 XX This sequence encodes a *Borrelia burgdorferi* (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the *Borrelia* genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the *Borrelia* genus. The products can also
 CC be used for detection of members of the *Borrelia* genus.
 SQ Sequence 477 BP; 132 A; 62 C; 114 G; 169 T; 0 other:
 Query Match 42.9%; Score 369.4; DB 20; Length 477;
 Best Local Similarity 86.1%; Pred. No. 7.5e-58;
 Matches 409; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 281 TGAATGTTTAAAGCGCTTTTATTTGGAGCGGAGAAAACTTTAGCTTATGAAC 340
 DB 3 TGAATGTTTAAAGCGCTTTTATTTGGAGCGGAGAAAACTTTAGCTTATGAAC 62
 QY 341 TGTATGAAGAATCCATTTGTCATTTTATTTGAACCTTTTATTTAGGCTTGAATAG 400
 DB 63 TAGCAAGCAAGATCTTATTTGACATTTTATTTGAACCTTTTATTTAGGCTTGAATAG 122
 QY 401 TTTCTTTGCTCAAGAGATATCTTGGGGGCTTTTCTTATTTCTTGATTTGATGACAGTGG 460
 DB 123 CTCCTTTGCTCAAGAGATATCTTGGGGGCTTTTCTTATTTCTTGATTTGATGACAGTGG 182
 QY 461 TATGAGGTTAATCTTACAGAGCTTATTTAGATCAAAAGCTTTGATPAAGATGCTCC 520
 DB 183 TATGAGGTTAATCTTACAGAGCTTATTTAGATCAAAAGCTTTGATPAAGATGCTCC 242
 QY 521 AAAAGCGCTTTTAAAGTGGCTTGGGTAAGGAATGATGGGCGAGCTTACTAT 580
 DB 243 AAAAGCTGCTTTTAAAGTGGCTTGGGTAAGGAATGATGGGCGAGCTTACTAT 302
 QY 581 GGCTGTGACAGATTTGACAGAAATTTATTCGTTTACATTTGCTAATAGTTAATAG 640
 DB 303 GGCTGTGACAGATTTGACAGAAATTTATTCGTTTACATTTGCTAATAGTTAATAG 362
 QY 641 GAAACTGAAAAATAGCTTAAATATAGCTTTGGAGGCTTGGACCTAGTTGATATTA 700
 DB 363 GAACTGAAAAATAGCTTAAATATAGCTTTGGAGGCTTGGACCTAGTTGATATTA 422
 QY 701 TAGGGCCCAAGCTAGGCTTGGGTTTGAACATCTTCAAAAAAGTTATTA 755
 DB 423 AATGGCCCAATCAGTGTCTTGGGTTTGAACATCTTCAAAAAAGTTATTA 477
 RESULT 6
 AAX20266
 ID AAX20266 standard; DNA; 7074 BP.
 XX
 XX AAX20266;
 AC
 XX 04-MAY-1999 (first entry)
 DT
 XX
 DE *Borrelia burgdorferi* polynucleotide sequence #19.
 XX
 KW *Borrelia burgdorferi*; spirochete; bacterium; pathogen; Lyme disease;

KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
 KW infection; diagnosis; characterisation; detection; ds.
 XX
 XX *Borrelia burgdorferi*.
 OS
 PN WO9858943-A1.
 PD
 XX 30-DEC-1998.
 PF
 XX 18-JUN-1998; 98WO-US12764.
 PR
 XX 03-SEP-1997; 97US-0057483.
 PR 20-JUN-1997; 97US-0050359.
 PR 22-JUL-1997; 97US-0053344.
 PR 22-JUL-1997; 97US-0053377.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMUNE INC.
 XX Claydon R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
 PI White OR;
 XX WPI: 1999-081217/07.
 XX
 XX New isolated *Borrelia burgdorferi* nucleic acids - used to develop
 PT products for the detection, diagnosis, characterisation, prevention
 PT and therapy of infections, particularly Lyme disease
 XX
 XX Claim 1: Page 951-955; 1128pp; English.
 XX
 XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
 CC *Borrelia burgdorferi* (Bb). Products derived from Bb can be used for
 CC the detection, diagnosis, characterisation, prevention and therapy of
 CC Bb infections, e.g. Lyme disease. They can also be used for the
 CC production of biosynthetic products, e.g. enzymes. *Borrelia* belongs
 CC to a family of motile, spiral-shaped bacteria called spirochetes.
 CC Spirochetes are pathogenic in humans and *Borrelia* causes epidemic and
 CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
 CC Lyme disease.
 SQ Sequence 7074 BP; 3064 A; 906 C; 866 G; 2233 T; 5 other:
 Query Match 11.6%; Score 100; DB 20; Length 7074;
 Best Local Similarity 50.5%; Pred. No. 2.6e-09;
 Matches 410; Conservative 0; Mismatches 355; Indels 47; Gaps 5;
 QY 39 CAATTAAGACATTTAATAAGATTAAGTAATTTATTTATGTCATAAGAGCT 98
 DB 2982 CTAAATAGAGATCTCTATTTGATTCATCAAAAAGTATTTAAATCAAAATCAACCTAT 3041
 QY 99 CTATGGGAAGGAATTTCAAGTAATCTTGAAGAAAGTTAATTAATAGT-TTAA 157
 DB 3042 TTATTAACAAATCATCTTAACCCAAATTTATAGAAATTAATAGTATCTTGG 3101
 QY 158 AAACCTTTTAAATTTCAATTAATATGTTACTTAATACAGTTTAAATAGAGGTTT 217
 DB 3102 GGGTTTTGATAAATTCATTAATATTAATTAATTAATTTATTTGGAGTAATCA 3161
 QY 218 TATGAATTAATTTAATTTGTTGTTGCTAGCCTTTGTTGTTTCTTACCTTTGCTCA 277
 DB 3162 TATGAATTAATTTAATTTGTTGTTGCTAGCCTTTGTTGTTTCTTACCTTTGCTCA 3221
 QY 278 AGCTGATGATTTCAAAAGGCTTTTAAATTTGGAGGGAAGAAACCTTTAGCTTATCA 337
 DB 3222 ACAAGTAAAGCTTGAAGAAAGTGT- - - - -GGAAGTATTTGAACCATTAATGAATTA 3275
 QY 338 AACTAGTAAGAAAGATCTATTTGCTATTTATTTATTTAGGCTTTGAAT 397
 DB 3276 AAGCGAAAAAGCACTATCTAGACCATCTTTTGAATTTATTTTAACTTTAGGAAT 3335
 QY 398 AGGTTCTTTTGTCAAGAGATATTTCTTGGGGGTTTCTTATTTCTTGATTTGATGACAGT 457
 DB 3336 AGGATCTTTTGTCAAGAGATATTTCTTGGGGGTTTCTTATTTCTTGATTTGATGACAGT 3395


```

Db 301 GCAAGATCCTATTCACATTTTATGACCTTTTATGAGGTTGGAATAGGCTCCT 360
Qy 407 TGTCTAAGAGATATTTTGGGGGTTTCTTATTTCTTGATTTGATGACGTTGATAGG 466
Db 361 TGTCTAAGAGATATTTTGGGGGTTTCTTATTTCTTGATTTGATGACGTTGATAGG 420
Qy 467 GTTATATCTACAGAGCTTATTTATGATATCAAGCTTGTGATAGAGTCCAAAAGC 526
Db 421 GCTTATATCTGCGGGGCTTATTTGATATCAAGGCTTGTGATAGTATTAATAAGC 480
Qy 527 CGCTTTAAGTACTGGGGTGAAGGATGATGTTGGCAGTGCATGCTATGCTGT 586
Db 481 TGTCTTCAAGTACTGGGGTGAAGGATGATGTTGGCAGTGCATGCTATGCTGT 540
Qy 587 GACAAGATGACAGAAATTTATTTCCGTTTACATTTGCTATAGTATTAATAGAACT 646
Db 541 GACAAGATTAACAGAAATTTATTTCCATTTTACATTTGCTATTAATAGAACT 600
Qy 647 GAAAAATAGCCTTATATAGCTTTTGAAGGTTGAGCCTAGTTTGTATTAATATAGG 706
Db 601 AAAAATAGCCTTATATAGCTTTTGAAGGATTTGAACTAGTTTGTATTAATAGG 660
Qy 707 CCAAGCTAGCCTCTTGGGTTTGAACATTTCTTCAAAAAAGTTATTAATTTAT 766
Db 661 CCAATCCAGTCTCTTGGGTTTGAACGTTCTTCAAAAAAGCTTATTAATTTA-TTAT 719
Qy 767 TATTAATGAGTGTATGACATTTTGTATTTGATGCT 805
Db 720 TACAAAAATGGGTATTCGAATTCGTATGAATGGT 758

```

RESULT 4
AAx61505
ID AAX61505 standard; DNA; 540 BP.

AC AAX61505;

DT 19-JUL-1999 (first entry)

DE B. burgdorferi antigenic protein coding sequence, f933.nt.

KM Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

OS Borrelia burgdorferi.

XX MO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;

XX WPI; 1999-189980/16.

XX P-PSDB; AAY19808.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

XX products for the diagnosis, prevention and treatment of diseases

XX caused by Borrelia, particularly Lyme disease

XX Claim 1; Page 73; 275pp; English.

XX This sequence encodes a Borrelia burgdorferi (Bb) protein of the

XX invention, which is suitable for use in a vaccine. The Bb polypeptides

CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

SO Sequence 540 BP; 145 A; 71 C; 123 G; 201 T; 0 other;
Query Match 48.7%; Score 420; DB 20; Length 540;
Best Local Similarity 87.4%; Pred. No. 5.9e-67;
Matches 472; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

```

Qy 219 ATGATAAATTTTAAATTTGTTTGGTACCTTTTGTGTTTCTTGTAGCTTGCCTCA 278
Db 1 ATGATAAATTTTAAATTTTGTGTTTGGTACCTTTTGTGTTTCTTGTAGCTTGCCTCA 60
Qy 279 CCGTATGTTCTTAAA--GCGCTTTTATTTGGGAGGAGGAGAAACTTTAGCTTAT 335
Db 61 GCTATGATTTCTTAAAATGAGGCGTTGGATGAGTCTGAGAGAAACTTTTGGTTTAT 120
Qy 336 GAACTAGTAAAGAAATCTATTTGCTATTTTATTAATGAACTTTTGTAGGTTTGA 395
Db 121 GAACTAGTAAAGAAATCTATTTGCTATTTTATTAATGAACTTTTGTAGGTTTGA 180
Qy 396 ATAGCTTTTGTCTCAAGAGATATCTTGGGGTTTCTTATTTCTGATTTGATGCA 455
Db 181 ATAGCTCTTTGCTCAAGAGATATCTTGGAGTTCTTATTTCTGATTTGATGCG 240
Qy 456 GTTGTATAGGCTTATATCTTACAGAGCTTATTTAGATATCAAGCTCTTGTATGAT 515
Db 241 GTTGTATAGGCTTATATCTTACAGAGCTTATTTAGATATCAAGCTCTTGTATGAT 300
Qy 516 GCTCAAAAGCGCTTTAAGTGGAGCTTGGGTTGAAGGATGATGTTGGAGTGCAT 575
Db 301 ACTAAAAAGCTCTTTTCAATGAGCTTGGGTTAAGGAGTTATGACAGGTGTGCTT 360
Qy 576 ACTATGCTGTGACAAAGATGACAAATTTATTTCCGTTTACATTTGCTAATAGTTAT 635
Db 361 ACTATGCTGTGACAAAGATGACAAATTTATTTCCATTTTACATTTGCTAATAGTTAT 420
Qy 636 AATAGGAACGTGAAAAATAGCTTATATTTAGCTTTTGGAGGTTGAGCCTAGTTTAT 695
Db 421 AATAGGAACGTGAAAAATAGCTTATATTTAGCTTTTGGAGGATTTGAACTAGTTTAT 480
Qy 696 ATTAATATGGGCAAGCTAGGCTTGTGGTTTGAACATCTTCAAAAAAGTTATTA 755
Db 481 GTTGCAATGGGCAATCCAGTGTCTTGTGTTGAACTGTCTTCAAAAAAGCTATTA 540

```

RESULT 5
AAx61506
ID AAX61506 standard; DNA; 477 BP.

AC AAX61506;

DT 19-JUL-1999 (first entry)

DE B. burgdorferi antigenic protein coding sequence, t933.nt.

KM Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

OS Borrelia burgdorferi.

XX MO9859071-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12718.

XX 03-SEP-1997; 97US-0057483.

XX 20-JUN-1997; 97US-0050359.

XX 22-JUL-1997; 97US-0053344.

XX 22-JUL-1997; 97US-0053377.

CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.
 XX
 XX
 SQ Sequence 749 BP; 234 A; 82 C; 150 G; 283 T; 0 other;

Query Match 70.4%; Score 607; DB 20; Length 749;
 Best Local Similarity 90.1%; Pred. No. 1.3e-100;
 Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

```

QY 27 TAGAATTTTCGCAATTAAGACATATTAAGAAATGAAATGCAATTTATGTCGA 86
DB 1 TGAATTTTCACAAATTAAGATATGTTAAAGAAATGAAATGCAATTTATGTCGA 60
QY 87 AATCAAGAAGCTCTATTTGGGAAGCAATTTCAAGTAATCTTGAAGAAAAAGTAAATTA 146
DB 61 AATCAAGAAGCTCTATTTGGGAAGCAATTTCAAGTAATCTTGAAGAAAAAGTAAATTA 120
QY 147 AATAGTTTAAACCTTTTAAATTTCAATTAATATGTTACTATAATACCAAGTTTAAAT 206
DB 121 AATAGTTTCAAAAGCCCTTTTAAATTTCAATTAATATGTTACTATAATACCAAG-TTAAAT 179
QY 207 AAGAGGTTTTTATGATTAATTTTAAATTTGTTTGTACCTTTTGTCTTTTCTTCT 266
DB 180 AAGAGGTTTTTATGATTAATTTTAAATTTTAAATTTTGTATTCCTTTTGTCTTTTCT 239
QY 267 AGCTTTGCTCAAGCTGATTTCTAAAGCGCTTTTAAATTTGGAGCGGAGAAAAACTT 326
DB 240 AGTTTTCGTCGA---GATGTTCTAAAGACTTTTAATCTGGAGCGGAGAAAAATTT 296
QY 327 TTAGCTTAAGAACTATGTAAGAAAGATCCATTTGTCATTTTAAATTTAAATTTTAA 386
DB 297 TTGCTTAAAGAACTATGTAAGAAAGATCCATTTGTCATTTTAAATTTTAAATTTTAA 356
QY 387 GGGTTTGAATAGGTTCTTTTGTCTCAAGGATATCTTGGGGGTTTTCTTATCTTGA 446
DB 357 GGGTTTGGATAGGTTCTTTTGTCTCAAGGATATCTTGGAGGTTCTTATCTTGA 416
QY 447 TTTGATGCAATGCTTATGAGGTTAATCTTACAGAGCTTATTTAATATCAAGCTCTT 506
DB 417 TTTGATGCGGTTGCTTATGAGGTTAATCTTACAGAGCTTATTTGACATCAAGATTTT 476
QY 507 GATAGAAATGCTCAAAAGCGCTTTTAAAGTGACCTTGGGTAAGGAATGATGTTGCA 566
DB 477 GATAGAAATGCTCAAAAGCGCTTTTAAAGTGACCTTGGGTAAGGAATGATGTTGCA 536
QY 567 GGTGCAATGCTTATGCTGCTGACAAATGTAAGTAATTTCTTCCGTTTAAATTTGCT 626
DB 537 GGTGCTTACTATGCTGCTGACAAATGTAAGTAATTTCTTCCATTTTAACTTTGCT 596
QY 627 AATAGTTTAAATGGAAGCTGAAATTAATAGCTTATAGCTTTTGAGGCTTGAAGCT 686
DB 597 AATAGTTTAAATGGAAGCTGAAATTAATAGCTTATAGCTTTTGAGGCTTGAAGCT 656
QY 687 AGTTTGAATTAATTAATGAGGCAAGCTAGCGCTTTGGTTTGAATCTTTCAAAAA 746
DB 657 AGTTTGAATTAATTAATGAGGCAAGCTAGCGCTTTGGTTTGAATCTTTCAAAAA 716
QY 747 AGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 773
DB 717 AGCTATTAATTTTATTTATCTAGAAAA 743

```

RESULT 3
 AAX30098
 ID AAX30098 standard; DNA; 759 BP.

AC AAX30098;
 XX
 DT 17-JUN-1999 (first entry)
 XX
 DE Borrelia burgdorferi B31 protein encoding cDNA.
 XX
 KW Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.

```

XX OS Borrelia burgdorferi.
XX FH Key Location/Qualifiers
XX CDS 170..709
XX FT /tag=a
XX PN WO9912960-A2.
XX PD 18-MAR-1999.
XX PF 04-SEP-1998; 98WO-1B01424.
XX PR 16-SEP-1997; 97US-0059036.
XX PR 10-SEP-1997; 97DK-0001041.
XX PA (SYMB-) SYMBICOM AB.
XX PI Bergstroem S;
XX DR WPI: 1999-215027/18.
XX DR P-PSDB; AA104278.
XX PT Nucleic acid from Borrelia burgdorferi encoding virulence associated
XX PT protein P13
XX PS Claim 1: Page 107-108; 118pp; English.
XX CC The present sequence encodes a Borrelia burgdorferi B31 protein. The
XX CC present invention describes an isolated nucleic acid (I) that: (i)
XX CC encodes a polypeptide fragment (II) immunologically reactive with
XX CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
XX CC Borrelia burgdorferi B31, but not reactive with most proteins from
XX CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
XX CC to specific nucleic acid sequences but not to genomic DNA from most
XX CC other spirochaetes. (I), (II) and transformed cells are useful in
XX CC vaccines to protect against infection by B. burgdorferi sensu lato.
XX CC (I), (II) and antibodies raised against (II) are used to diagnose such
XX CC infections. In standard immunoassays or amplification/hybridization
XX CC tests, (I) are also used to produce recombinant (II). The 13 kD protein
XX CC is involved in virulence and is highly conserved within B. burgdorferi
XX CC sensu lato, but is absent from other Borrelia species (e.g. those
XX CC responsible for relapsing fever or avian borreliosis). It should provide
XX CC a more specific and sensitive antibody response, and diagnosis,
XX CC compared with use of whole bacteria as antigen.
XX SQ Sequence 759 BP; 228 A; 90 C; 157 G; 284 T; 0 other;
XX
XX Query Match 69.2%; Score 596.6; DB 20; Length 759;
XX Best Local Similarity 88.4%; Pred. No. 1e-98;
XX Matches 671; Conservative 0; Mismatches 84; Indels 4; Gaps 2;
XX
QY 50 ATTATTAAGAAATGAAATTTGCTAATTTTATGTCAAATCAAGAGCTCTATGGGAAG 109
DB 1 ATTGTTAAAGAAATGAAATTTGATTAATTTATGTCAAATCAAGAGCTCTATGGGAAG 60
QY 110 CGAATTCAGTAATTAATCTTTGAAAAAGTAAATTAATTAATAGTTTAAAAACCTTTTAA 169
DB 61 CGAATTCAGTAATTAATCTTTGAAAAAGTAAATTAATTAATTAATTAATTAATTAATTA 120
QY 170 AATTCATTAATTAATGTTACTATACAGCTTTTAAATTAAGAGGTTTATGAAATTAAT 229
DB 121 AATTCATTAATTAATGTTACTATACAGCTTTTAAATTAAGAGGTTTATGAAATTAAT 180
QY 230 TTTAATTTGTTTGTCTAGCCCTTTTGTGTTTCTAGCTTTGCTCAAGCTGATGATTC 289
DB 181 TTTAATTTGTTTGTGGAACCTTTTGTGTTTCTAGCTTTGCTCAAGCTGATGATTC 240
QY 290 TAAAG--GCGCTTTTAATTTGGAGCGGAGAAAAACTTTTATGCTTAAGAACTGTAA 346
DB 241 TAAAGATGCGCTTTGAGGATGAGTGTGAGAAAAACTTTTATGTAAGAACTGTAA 300
QY 347 GAAAGATCCATTTGTCATTTTATTTGAACCTTTTATGAGGTTTGAATAGTTCTTT 406

```

DR P-PSDB: AAY04279.
 XX Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
 PT protein p13
 PS Claim 1; Page 110-111; 118pp; English.
 XX The present sequence encodes a *Borrelia afzelii* ACAI protein. The
 CC present invention describes an isolated nucleic acid (I) that: (1)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (II) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within *B. burgdorferi*
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.
 XX
 XX
 SQ Sequence 862 BP; 265 A; 94 C; 162 G; 341 T; 0 other;
 Query Match 100.0%; Score 862; DB 20; Length 862;
 Best Local Similarity 100.0%; Pred. No. 1.8e-146;
 Matches 862; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GATTTTCATTGATCCGAGATTGTGATTTTCGACAAATAGACATTATTAAG 60
 Db 1 GATTTTCATTGATCCGAGATTGTGATTTTCGACAAATAGACATTATTAAG 60
 Oy 61 AATTGAAATTCCTAATTTATGTCGAATCAAGAACTCTATTTGGGAGCGAATTCAAG 120
 Db 61 AATTGAAATTCCTAATTTATGTCGAATCAAGAACTCTATTTGGGAGCGAATTCAAG 120
 Oy 121 TTAATCTTGAAGAAAGTTAATTAATTAATTAATTAATTAATTAATTAATTA 180
 Db 121 TTAATCTTGAAGAAAGTTAATTAATTAATTAATTAATTAATTAATTAATTA 180
 Oy 181 TATGTTACTATTAATACAGATTTAAAGAGGTTTATATGAAATTTTAAATTTGTTG 240
 Db 181 TATGTTACTATTAATACAGATTTAAAGAGGTTTATATGAAATTTTAAATTTGTTG 240
 Oy 241 TTTTCTAGCCTTTGTGTTTCTTACCTTTGCTCAAGCTGATCTTAAACGCTT 300
 Db 241 TTTTCTAGCCTTTGTGTTTCTTACCTTTGCTCAAGCTGATCTTAAACGCTT 300
 Oy 301 TTAATTTGGAGCGGAGAGAAACCTTTAGCTTATGAACATGTAAGAAAGTCTATTG 360
 Db 301 TTAATTTGGAGCGGAGAGAAACCTTTAGCTTATGAACATGTAAGAAAGTCTATTG 360
 Oy 361 TGCCATTTTATGTAACCTTTTATAGGTTTGAATAGGTTCTTTGCTCAAGAGATA 420
 Db 361 TGCCATTTTATGTAACCTTTTATAGGTTTGAATAGGTTCTTTGCTCAAGAGATA 420
 Oy 421 TTTCTGGGGGTTTCTTATTTCTGATTTTGATGAGTTGGTATAGGTTATATTACAG 480
 Db 421 TTTCTGGGGGTTTCTTATTTCTGATTTTGATGAGTTGGTATAGGTTATATTACAG 480
 Oy 481 GAGCTTATTTAGATTCGAAGCTCTTGATAGAAATGCTCAAAACCCGTTTAAAGTGA 540
 Db 481 GAGCTTATTTAGATTCGAAGCTCTTGATAGAAATGCTCAAAACCCGTTTAAAGTGA 540
 Oy 541 CTGGGGTAAGGAATGATGTTGGCAGGTGCAATTACTATGCTTGACAAAGATTGACAG 600
 Db 541 CTGGGGTAAGGAATGATGTTGGCAGGTGCAATTACTATGCTTGACAAAGATTGACAG 600
 Oy 601 AATATATATTCGGTTACATTTGCTATAGTATATAGGAAATGAAATATACCTTA 660
 Db 601 AATATATATTCGGTTACATTTGCTATAGTATATAGGAAATGAAATATACCTTA 660

Db 601 AATATATATTCGGTTACATTTGCTATAGTATATAGGAAATGAAATATACCTTA 660
 Oy 661 ATATAGCTTTTGGAGGGTTGAGCCTAGTTTGAATATATATATGAGCCAACTAGCCTC 720
 Db 661 ATATAGCTTTTGGAGGGTTGAGCCTAGTTTGAATATATATATGAGCCAACTAGCCTC 720
 Oy 721 TTGGGTTGAACATCTTTCAAAAAAGTTATATTTTATTTATTTATTTAAATGAGTGC 780
 Db 721 TTGGGTTGAACATCTTTCAAAAAAGTTATATTTTATTTATTTATTTAAATGAGTGC 780
 Oy 781 ATAGCAATTTTGTATTTGATTTGCTCATTTGTAATTTGAAATATAGCTTTGTTATAT 840
 Db 781 ATAGCAATTTTGTATTTGATTTGCTCATTTGTAATTTGAAATATAGCTTTGTTATAT 840
 Oy 841 TTAATTTTATTTCTGCTGTA 862
 Db 841 TTAATTTTATTTCTGCTGTA 862
 RESULT 2
 ID AAX30100 standard; DNA; 749 BP.
 AC AAX30100;
 DT 17-JUN-1999 (first entry)
 XX
 DE *Borrelia garinii* IP90 protein encoding cDNA.
 XX
 KW *Borrelia*; p13 antigen; spirochaete; vaccine; infection; diagnosis; ss.
 XX
 OS *Borrelia garinii*.
 XX
 FH Key location/Qualifiers
 FT CDS 192..725
 XX /*tag= a
 PN W09912960-A2.
 XX
 PD 18-MAR-1999.
 XX
 PE 04-SEP-1998; 98WO-IB01424.
 XX
 PR 16-SEP-1997; 97US-0059036.
 XX
 PR 10-SEP-1997; 97DK-0001041.
 XX
 PA (SYMB-) SYMBICOM AB.
 XX
 PI Bergstroem S;
 XX
 DR WPI: 1999-215027/18.
 XX
 DR P-PSDB: AAY04280.
 XX
 PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
 PS protein p13
 PS Claim 1; Page 113-114; 118pp; English.
 XX The present sequence encodes a *Borrelia garinii* IP90 protein. The
 CC present invention describes an isolated nucleic acid (I) that: (1)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (II) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within *B. burgdorferi*
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:05:42 ; Search time 190.586 Seconds
(without alignments)
10185.565 Million cell updates/sec

Title: US-09-508-487-20

Perfect score: 862
Sequence: 1 gattttcattgcgatccag.....atatattctctcgtctaa 862

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
N.Geneseq_101002: *
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: *
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: *
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	862	100.0	862	20	AAAX30099
2	607	70.4	749	20	AAAX30100
3	596.6	69.2	759	20	AAAX30098
4	420	48.7	540	20	AAAX61505
5	369.4	42.9	477	20	AAAX61506
6	100	11.6	7074	20	AAAX20266
7	94.4	11.0	1761	20	AAAX20310
8	88.2	10.2	9399	20	AAAX20261
9	67.4	7.8	14950	24	ABL33257

10	67	7.8	5930	24	ABL32517	Human immune syste
11	67	7.8	6012	22	AA546298	Tumour suppressor
12	67	7.8	6012	24	ABL70148	Chemically treated
13	67	7.8	6012	24	ABK31187	Signal transductio
14	65.8	7.6	9742	24	ABL70480	Chemically treated
15	65.2	7.6	11036	22	AA545410	Chemically pretrea
16	65.2	7.6	11036	24	ABK28263	DNA transcription
17	64.4	7.5	6590	24	ABL70298	Chemically treated
18	64.4	7.5	6590	24	ABL33383	Human immune syste
19	64.4	7.5	6590	24	AA561228	Human immune syste
20	64.4	7.5	19087	24	ABL32793	Human immune syste
21	63.8	7.4	6652	24	ABQ66988	Human anglogenesis
22	63.8	7.4	17967	24	ABK63015	Human immune syste
23	63.4	7.4	5815	24	ABL33235	Human immune syste
24	63.4	7.4	16766	24	ABL34157	Human immune syste
25	63.2	7.3	5413	22	AA546694	Tumour suppressor
26	63.2	7.3	7441	24	ABK40058	Human chemically p
27	63.2	7.3	7479	24	AA563345	Human chemically p
28	63.2	7.3	13784	24	ABK40062	Human gene regulat
29	63.2	7.3	19380	24	AA561426	Human gene regulat
30	62.6	7.3	6716	24	ABL33782	Human immune syste
31	62.4	7.2	5963	24	ABL34613	Human metastasis a
32	62	7.2	6509	24	AA561086	Human gene regulat
33	62	7.2	6509	24	ABK31189	Signal transductio
34	61.8	7.2	5884	24	ABL34165	Human immune syste
35	61.4	7.1	14987	24	ABL32631	Human immune syste
36	61.4	7.1	15121	24	ABN80238	Human chemically m
37	61.4	7.1	73334	24	ABL92319	Human immune syste
38	61.4	7.1	73334	24	ABL34125	Human immune syste
39	60.8	7.1	11745	24	ABK28332	DNA transcription
40	60.8	7.1	53585	20	AAAX20251	Borrelia burgdorfe
41	60.6	7.0	15825	24	ABL33242	Human immune syste
42	60.6	7.0	15825	24	ABL34558	Human metastasis a
43	60.4	7.0	6244	24	ABL32484	Human immune syste
44	60.4	7.0	7771	24	ABL33973	Human immune syste
45	60.2	7.0	6298	24	ABL34470	Human metastasis a

ALIGNMENTS

RESULT 1	
AAAX30099	
ID	AAAX30099 standard; DNA; 862 BP.
XX	
AC	AAAX30099;
XX	
DT	17-JUN-1999 (first entry)
XX	
DE	Borrelia afzelii ACAI protein encoding cDNA.
XX	
KW	Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.
XX	
OS	Borrelia afzelii.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
XX	/tag= a
XX	
PN	WO9912960-A2.
XX	
PD	18-MAR-1999.
XX	
PF	04-SEP-1998; 98WO-IB01424.
XX	
PR	16-SEP-1997; 97US-0059036.
XX	
PR	10-SEP-1997; 97DK-0001041.
XX	
PA	(SYMB-) SYMBICOM AB.
XX	
PI	Bergstroem S;
XX	
DR	WPI; 1999-215027/18.

THIS PAGE BLANK (USPTO)

CDs

```

/product="Vegetative specific protein H7"
/protein_id="AAM43693.1"
/db_xref="gi:21166076"
/translation="MDVOTKYGAGONKVLGAGNKKIYSEEDIALPELNPVPOAIO
RARNALKTQELAKINERPGVINEYSSGSAIPSOAVLSKLEKLNKLN"
join(17001..17789,17899..18159)
/notes="ORF_ID:dd_02308"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM43694.1"
/db_xref="gi:21166077"
/translation="MSIYIDNINNNNNKLIINYLINNDCLVMDVGDILIRKRYL
LGSIVGSENLRLNSTINSVPLISPEIRLGEKQWQIIESAEPFLIDHNSLE
KESRQTYOSLMDKKEORLKEKKEFRKLNNTSNKKNKIPSLDPTDNNNN
TNNEENNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNKEEIK
EKEEINSSKNIFFVYDISTSCTKEADINERIPYTLWRDGDYIILSRGYLSSGS
KFGTEFLAKGEPFLVATFTVIRKSNQTSQSDITITARLAVHVNNTLLASLDEE
TNKPIYLIINMKVGS"
join(19303..20698,21044..21077,21937..22371,22484..24078,
24154..26522)
/notes="ORF_ID:dd_03155"
/codon_start=1
/product="Opaque-specific ABC transporter CDR3."
/protein_id="AAM43695.1"
/db_xref="gi:21166078"
/translation="MSIISSEFISIPSTINGKDKSINLNNDNNKNNNNNSIKN
NNNNKNNPSDNDIKPREGGIDDKLTITATSLTPPKGDIIESDPLSKLITDIT
NDNSSTTVCGCDDNEEOOQOQOEOQOEOQOHEFKIPMIOIOTPOINDNIE
WPTWKDEENRDLNRLISRLDTSNRLNGSEKENVYQOTKFSKLYMGKNSPIF
GDIKEDPNNNNFDTTASPTSGGDKOLIRSYNGINVEDRMLGIVYMGKNSPIF
PGCVGALASMFDSMACIRFEYNNNDLNLKOKOKOQOQOEOPOGNGRI
VLPITNSKTYVTANLSVYKKEVLSSEFSSEVERIKORDEKKEFISITLDKQNGVI
RAESTAIWYKIDPITSTSSPSSELSPLSPSTNENKNNNNNNNNNNNNNNNN
NNNNNNIOPDNNSSCKCHSYDTIDMLSENNNNNNNNNNNNNNNNNNNNNNNN
DYSIDNINTEEFENVSKLEQOSIKFEIDGKNNNNNDIEGKRPENEDFKLQY
FEDSQOKMLINHKPKMGVSKLNLTVGQADNSVIDNSPFPALKILNPFYFK
KDKNINTEFNILNDINAFIEDGKMLVLRGAGCSTLRLVLANRESYISVDGVTY
GNIAASDKYRGTELTGVEEDIHPTLVEEDLDEFTLTKLFGSRLEESKRNRTK
IYELLSVNGYLVNGDITVGNDFVGRSLGSGEKRITITTEAVSGSLISCMQSTGID
AASAVDYAKSLRIMSDPTDKTTIASFYAASSIYVLPKVIYLDGRCRITRPIELAK
OYFLDGFDCPRKSTPPFLGITNPRIVKGVGPEVSVLDLPAWKMSQLFOS
MKHOLEKQVEQKPSVDFREOVLPNKRSTRNSSEYSAQTLAQLORSLT
WGDRTTSLRLLIYLSIFIGGAIYFOQLPTDGTGELGALFISIIIFNCILTOELH
GALSGRILQKHSYALYRPSAYFSQILIDIFELVYVLFSLFVYMGYEPYADK
FEIFCELTGVSSLSASLFRGFANFTPSLFTQONLNVFIEVYFAYSOCPDMHS
WEKTTYVINPLAFAKSLMINEFGKIDSCDSANLPRHPNNSTYSMDSHRCAPGS
IEGSLSVGENYIMDALDINSRHLNLYVFLPEFLITGLNFAVEYEDTSGGITH
KYVKGKAPKLNDEEERNNOIVAKATDNKMDTLKMRGLESMSISYVPAVATNK
LLLDIMGIRPGQWNTALMGSGAGTLLDYLAKRMGTGVTGSLNGKOLEIDFE
RITGVEMDVHNPGLIVREALREFSAGLRQEWYPLKDYQYEHVLEMEKHLGDA
LIGLETGVGISVEERKRLTIGVELVAPOLIFLDEPTSGLDAOSSYNIKIFIRLAD
AGMPLVCIHQSSVLPFHPDRILLARGKTVYRGDIDGSKTITTSFERGVAPCT
ESENPAEYILBATGAGTIGKTDVNPPEAKROSSYONVNEIDLRTKEDEKYLIDS
DLQVKGKAPREFRANGFLTOFIEVYKRLNIIYRDVEYTWGSAVSGLVIGFTF
YDLKNSSDOQORIFMSWEAMILGVLLIYLVLPMEFIOKEFEKRDTSKYSWHAESL

```

CDs

```

Query Match      7.7%: Score 66.4; DB 2: Length 129360;
Best Local Similarity 46.8%; Pred. No. 0.16;
Matches 383; Conservative 0; Mismatches 421; Indels 14; Gaps 5;

```

```

OY 38 ACAATTAAGACATTTAAAGAAATGTAATTTATGCTCAATCAAGAGAC 97
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18171 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTAATTAATAA 18230
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 98 TCTATTGGAGAGCAATTTCAAGTAATACCTTGAAAAAGTTAAATTAAGTTTAA 157
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18231 TATATATATACCAATGATCAAAAAAACAATAAATAAATAAAAAACA 18290
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 158 - AAACCTTTTAAATTCATTAATGCTTACTATATACCAAGTTTAAAGAGTT 215
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18291 ATAAAAAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 18350
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 216 TTTATGAATAAATTTTAAATGTTGTTGCTAGCCTTTGCTTTTCTAGCTTGGCT 275
    | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 18351 AATTTAAAGGGTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 18410
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 276 CAAGCTGATGATTTCTAAAGCG-CTTTTAATTTGGAGCGGAGAAAACCTTTAGCTTA 334
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18411 TCTTATTTGTTTTTAAAGAGTAGATTAACAATAATATATTTAATAACCTTTATTAAGT 18470
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 335 TGAACCTAGTAAGAAGATCCATTTGGCATTTTATTTGAACCTTTTATTTAGGGTTGG 394
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18471 AGTATTTACTATTATTGAAGAGATTTATCCCATTTCAAGTATATTTATTTTATTTT 18528
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 395 AATAGGTTCTTTTGCCTCAAGAGATATTCCTGGGGCTTTTCCTATTCCTGATTTGATGC 454
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18529 - - - - - TTTTATTTTATTTGTTTTTATTTTATTTTATTTTATTTTATTTTATTTTTC 18583
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 455 AGTTGATTAAGGTTTATTTACTTACAGAGCTTATTTAGATATCAAGCTGTGATAGA 514
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18584 AATTCATTTAAGGGGAAAAAAGAGAGATATTTAAATATTAACCTAACTAAATTTGTA 18643
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 515 TGCTCAAAAAGCGCTTTTAAGTGAAGTGGGTAAGGAGATGATGTCGAGGTGCAGT 574
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18644 TATATTAAAAAAGGAAAAAAGGAAAAAATTTTGGTAAAGTTATTTTATTTATATGTA 18703
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 575 TACTATGCTGTGACAAGATTTGACAGAAATATTTATTCGTTTACATTTGCTAATAGTTA 634
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18704 AATACGATATTATATATATTTGAAGGCTTATCGAGA - GATTAATAAATAATATATATA 18761
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 635 TAATAGAAACGAAATAATAGCCTTAATATAGCTTTTGAAGGCTTTGAGCCTTACTTTTGA 694
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18762 AATATAAATAATTAATAGTTCATATTAACAAGAAATTTATTCATTCATGAGAC 18821
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 695 TATTAATATGAGCGCAAGCTAGCGCTCTGGGTTGAGTATCTTCAAAAAAGTATTA 754
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18822 AACACAGCGCGGCTTAGAGAAATTTCTTTTATTTTATTTTATTTTATTTTATTTT 18881
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 755 AATTTATTTTATTTAATAATGAGATAGCAATTTTGTATTTGATGCTCATTTGTAAT 814
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18882 AATTTATTTTATTTT - TATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 18939
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 815 TGAAATTTAGAGCTTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 852
    | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 18940 TAACATTTAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 18977
    | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 27, 2003, 03:11:15
Job time : 2281.68 secs

QY 338 AACTAGTAAAGATCCATTGTCACATTTTATGAAACCTTTTATGAGGTTTGGAT 397
 Db 4435 TATTGAAGTGTGATATATTTTATTTAGAAATATATTTTATTTTATTTTAT 4494
 QY 398 AGTTCCTTTGCTCAAGAGATTTCTGGGGTTTCTTATCTTGATTTGATGAGT 457
 Db 4495 GAATGTATATAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4554
 QY 458 TGGTATAGGTTATTTACTTACAGAGCTTATTTAGATTTCAACCTCTGATTAAGATGC 517
 Db 4555 TTTTATGCTTGAAGTTTGAAGATTAATATATTTTATTTTATTTTATTTTATTTTAT 4614
 QY 518 TCCAAAAGCCCTTTTAACTGAGCTTGGGCTAAGGAATGATGTGCGAGCTGAGTAC 577
 Db 4615 AAGATTTTATAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4674
 QY 578 TATGGCTGTGACAGATGACAGAAATTTATTCCTTACATTTGCTAAATAGTTATTA 637
 Db 4675 AATTTAT 4734
 QY 638 TAGGAATCTGAAATATACCTTATATATATATATATATATATATATATATATATAT 4734
 Db 4735 TATTAAT 4794
 QY 698 TAAATATGGCCAGCTAGCTGCTTGGCTTGAACCTATCTTCAAAAAAGTTATTAAT 757
 Db 4795 TATAAAT 4854
 QY 758 TTATTTTATTTTAAATAGATGATAGCAATTTTGTATTTGATTTGCATTTGATTA 817
 Db 4855 ATTAATGCTAT 4914
 QY 818 AAATAGAGCTTTTGTATATATATATATATATATATATATATATATATATATAT 862
 Db 4915 TTTTAACTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4959
 RESULT 12
 LOCUS AX251052 6012 bp DNA linear PAT 05-OCT-2001
 DEFINITION Sequence 20 from Patent WO0168912.
 ACCESSION AX251052
 VERSION AX251052.1 GI:15984475
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 location/Qualifiers
 1. 6012
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 1842 a 43 c 1153 g 2973 t 1 others
 ORIGIN
 Query Match 7.8%; Score 67; DB 6; Length 6012;
 Best Local Similarity 44.4%; Pred. No. 0.32;
 Matches 364; Conservative 0; Mismatches 445; Indels 10; Gaps 2;
 QY 51 TTATTAAGAAGATTAATTTGCTAATTTTATGCTCAATCAAGAAGCTCTATTTGGAGAC 110
 Db 3973 TAAATATATATTTTAAATGAGCAAAATGCTTAAATTAAGGGCTATATAGATTTTAA 4032
 QY 111 GAATTTCAAGTATCTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTAA 170
 Db 4033 TAACTTTGAATTTAATGAGTAGTATTTTAAATTTTAAATTTTAAATTTTAAATTTT 4092

QY 171 ATTTCATTAATATGTTACTATATATACAGTTTATATATAAGGTTTATGA-----A 223
 Db 4093 ATTTTATGTTTATATATTTAGCTTATATGATGTAATAGCTGGGTTTATATGCTTTTGA 4152
 QY 224 TAAATTTTAAATTTGTTGTTTGTAGACCTTTTGTGTTTCTTGTAGCTTGTCAAGCTGA 283
 Db 4153 TATTTTGTGTTTGTGTTTGTAGTATTTTATTTTATTTTATTTTATTTTATTTTAT 4212
 QY 284 TGATTTCAAAAGCCCTTTTAAATTTTGGAGCGGAGAAAACTTTTATAGCTTATGAACATG 343
 Db 4213 TATGAGTGTGTTGAGTGTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4272
 QY 344 TAAAGAAATCCATTTTGTCCATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 403
 Db 4273 TAGGAGACAGAAATGCGGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 4332
 QY 404 TTTTGCTCAAGGACATATTTCTGGGGTTTCTTATTTTGTGATTTTGTATGATTTGAT 463
 Db 4333 TTTTGTGGGGGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4392
 QY 464 AGGTTATTTACTTACAGAGCTTATTTATGATATCAAGCTCTGATTAAGATGCTCAAA 523
 Db 4393 TTAGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4452
 QY 524 AGCCGTTTAACTGAGCTTGGGCTAAGGAATGATTTGCGAGCTGAGTTACTATGCG 583
 Db 4453 TTTATAGATTTTATAGGAGAGGGTAAATGTTATAGTGTGTTTATTTTATTTTATTT 4512
 QY 584 TGTGACAGATTTGACAGAAATTTATTTCCGTTTACATTTGCTATTAATTTAAATGAGA 643
 Db 4513 ---GATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4569
 QY 644 ACTGAAATATAGCCCTTAATATATAGCTTTTGAAGGTTTGAAGCTAGCTTTTGAATAT 703
 Db 4570 TGTGATTTTATTTGTTTATTTATATATATATATATTTTGTATTAATATTTTAAAGAG 4629
 QY 704 GGCCCAAGCTAGCCCTTTGGGTTGAACATCTTTTCAAAAAAGTTATTTATTTATTT 763
 Db 4630 TAGGAAGTTTAAATTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4689
 QY 764 TATTAATTAATAGATGATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTT 823
 Db 4690 TTGTTGTTATTTAGCTTTTAAAGTTGTTTATTTATTTTGGGTATTTTATTTATTTAT 4749
 QY 824 GAGCTTTGTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 862
 Db 4750 TTATTTTGTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4788
 RESULT 13
 LOCUS AX344183 6012 bp DNA linear PAT 01-FEB-2002
 DEFINITION Sequence 30 from Patent WO0200926.
 ACCESSION AX344183
 VERSION AX344183.1 GI:18492071
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source
 location/Qualifiers
 1. 6012
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 1842 a 43 c 1153 g 2973 t 1 others
 ORIGIN

gene
CDS
/product="similar to PIR:A64689 percent identity: 35.78;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66053.1"
/db_xref="GI:2690010"
/translation="MEINLOSGLNNKNNKNNKIFISCSLVSPDPNRTYSSNR
GOIDSPGVKIMKHITLTONNAILTTSLKTDPSHAFTVKSLSIIPGYEPIY
QIDEPKIFVDSNDFYTLGLGTDLSHLYKVPITINIGKFSFLDETLQVSL
IEKVFYEETNELYNOTSRFINVSKEVYNTIGSWOTLSYLYRGRKREMPSEVH
LEKVIKTIHYTGIOKK"
2104..2495
/gene="BBG03"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact.; similar to
GP:2239281 percent identity: 63.64; identified by sequence
similarity; putative; conserved hypothetical protein,
authentic frameshift"
complement(2750..2857)
/gene="BBG04"
complement(2750..2857)
/gene="BBG04"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG04"
/protein_id="AAC66074.1"
/db_xref="GI:2690031"
/translation="MKQBAISIFDLEIAVHIKTLINLNIYNGV"
complement(2891..4056)
/gene="BBG05"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact.; similar to
GP:2228714 percent identity: 99.12; identified by sequence
similarity; putative; transposase-like protein, authentic
frameshift"
4208..5368
/gene="BBG06"
4208..5368
/gene="BBG06"
/note="similar to GB:U03641 PID:458216 percent identity:
66.76; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66054.1"
/db_xref="GI:2690011"
/translation="MKVFTFLKLCITIIYNIPRSSMINNSKKNPNHKLKLOOLIV
LSTLAYNSKYNKTKQNLICPBNKRRGQPTTIRMQNTYIKREKPYRTTY
YKHGINGETIYKLNYSKRECYLIQHFREKDYFKARVDYLDKRNKNSVD
LVECLNNNNIKERKLIQIEKYQVIRFKCNLSLKEILPLINDINDDELKIL
YIKRIEINLTNNTLVNSYFKRONKRLKLSINTOOOLEKNGNPEOLEISQKI
KENYKTHPIENHKKYKDLSHIKRLEKSEIKERKENOKCENKRTNIFNILLQK
KEKVIDKPLIKIYLNKKNLEYNKVENNNYVELLEIKKKKLOLKEV"
5378..5950
/gene="BBG07"
5378..5950
/gene="BBG07"
/note="similar to PID:1143398 PID:1174342 PID:1174350
PID:2182755 percent identity: 57.84; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11

gene
CDS
/product="conserved hypothetical protein"
/protein_id="AAC66055.1"
/db_xref="GI:2690012"
/translation="MONMAKSIQLVKPIVCSNKKDLFIKIKDNKTIYTRIMDI
YKGLNKKKKKKRIFSLRGLFNOSKVEEENFELADKFLGIYGYKRIKIFVKQ
VCGLEKSYSLSKTYMVRKKSIFCYFSLRILKKEVNNTYNAALSMFTTLER
OVIDFDKRTIPQKGFPIKAIKSLKHHIL"
5911..6678
/gene="BBG08"
5911..6678
/gene="BBG08"
/note="similar to PID:1174351 percent identity: 39.26;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="stage 0 sporulation protein J (spooJ)"
/protein_id="AAC66056.1"
/db_xref="GI:2690013"
/translation="MDRRKLLAKRHIIAASIKGVKSTSSISTITLSTKNTVLLY
DLDPQNAVTSYFTQDHPRLMELINIVSYSLIKHKTFKDVVISLKNLPISTEL
AKFSKGNQFRELMLRNAVNYLEDYDVIIDPSSLSSELDNALVADVILIPLE
RMAVENPLILNIQKLELNNMGKFAKIHIFASKVEIGRTVSTEISSLKELYLNF
IGEVHSEALKKIDVAIGPENENYKEYLRILEKI"
6737..7288
/gene="BBG09"
6737..7288
/gene="BBG09"
/note="similar to PID:1143400 percent identity: 43.68;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66057.1"
/db_xref="GI:2690014"
/translation="MWKRVYIINDRIYRNATYINTERDREKXELKLNELNIEDD
IRANKINMKILLETIRNKLITLDGYKKFEDIFDKIARQAVYIKIAXLIFGKLE
EIDTIENGIDKTLNLMKDKINKANILITPLRPLEQGECDPTKNNPKRANLLED
FYORKNEQLIKLEEKOKYS"
complement(7483..10779)
/gene="BBG10"
complement(7483..10779)
/gene="BBG10"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG10"
/protein_id="AAC66075.1"
/db_xref="GI:2690032"
/translation="MKLDEIILPLSMISISNOQLDAIETLAKIAEOKFSSLDLKS
LESAKSGSNLEKALINKSSQNAIKNFSLANSVNSKAGSMKSIGVLSVGGLG
SVNLANKTGAFNOMLAAPITIAVKAIOAGSTTIGIDGADALDERNEVSTF
SDMIGNEDIGSLAEMRPFDETLFTDAIANTKTMISGATASEVERIRMGGA
AGGSECELEKLABYRSVSSNNQVLELVALRDAGVDTITLIAEGLAGEALYKKA
SDGKIGFALNKALSKATSEGGKFTGNTAKAKTIAQAOQOTAKSEKFLPDIGKALE
PMIGIFEVKQFLWALVPLKVTATIAITTKGELVAALAGVWGMKMLERAIT
TFPVLYEGIVYIGKLELATYVSSGFLDFFKGFEBIIMVOKLIDMISKATYK
NLVPEWKAKEKNDISGESEPRDKFEDNATINIKKAEDYOKLODEIFNRQDIYN
KTKGARBOALNLEKTIENKOKFDEKSIJEDLOTDENKILVGEKSVNEFNNSYN
DPVNETOULLKESEEREIITKTLPHPOVSALOKLONEITENKKAPEYKKGSELT
NESNQVVALEKOVNEYEKALDLSFVEQAKALOKETTDLEWETMLPAKERSAK
KMAKSIQAMVKKFVDEHKSQPKRLNENRNTIKOYAAQDPTTSLVDSMIDGLNVR
NAFMIDQKFLNKTGSEISIEEFNLINGDVNMBEOLKMTQOMYSMTGLTAA
GAVFGPGEAAVAILNGLDIPVWGLKGOEKRIKAIKIKDEDELEEKSEVELK
LEDREDEIKMRKEKLSLDDEVYKTEIFELKOAQSKOISGEEOKRLHDVOTEKTK
KDIEFTQITKTEBAKLEVERKNTSKIEPRIRQAQAVDVYNAEDMMNLGKPDRLK
TQKILDELTKRIAKVSAGSIEIKLAGGARFVSNKPTYPNPGSVMSSEFGQPELVR
ITPAPIDENLRLEAKIIAETITKLOKTOSSVYNNFYNNFNGVDLAEKLVRLKSK
EHIMGFRAAE"
complement(10776..11015)
/gene="BBG11"
complement(10776..11015)
/gene="BBG11"

TITLE
 JOURNAL
 Submitted (08-DEC-1997) The Institute for Genomic Research, 9712
 Direct Submission
 Location/Qualifiers
 Medical Center Dr, Rockville, MD 20850, USA
 1. 27323

FEATURES
 source
 /organism="Borrelia burgdorferi"
 /db_xref="taxon:139"
 /plasmid="p28-4"
 174. 608
 /gene="BB101"
 174. 608
 /gene="BB101"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB101"
 /protein_id="AAC6211.1"
 /db_xref="GI:2690108"
 /translation="MENQYFNFLEFFIKIIFILYIMKSVKKEQEKIDHACRIIL
 LLETIFENILLENYSOKLTKYNNENKLPNSITMKKYLKQKEKIKIIAK
 YEKNDOSLIYKKLVLTLEKILKLELFYKELKQIKQNTT"
 736. 1104
 /gene="BB102"
 736. 1104
 /gene="BB102"
 /note="similar to GP:1655791 percent identity: 66.67;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AAC66189.1"
 /db_xref="GI:2690086"
 /translation="MDGVINDTVARMKKQIKFNKNTLITLVKTLIDHKNKELYSANK
 TYNVLIQNNNEALAKTYQLRVNKKILELLEKPNKVIILKRPYKKNESFGLY
 TLLSPLEVCCTKIYNLHPNT"
 complement(1847. 1972)
 /gene="BB103"
 complement(1847. 1972)
 /gene="BB103"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB103"
 /protein_id="AAC66212.1"
 /db_xref="GI:2690109"
 /translation="WVYVLDPNLSLYFGGLACWDFYVWVWISLSEIVSFL"
 complement(2124. 2219)
 /gene="BB104"
 complement(2124. 2219)
 /gene="BB104"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB104"
 /protein_id="AAC66213.1"
 /db_xref="GI:2690110"
 /translation="WGTIQKFLIKHGLMRAKPIGESKYFIY"
 complement(2188. 2310)
 /gene="BB105"
 complement(2188. 2310)
 /gene="BB105"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1

gene
 CDS
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB105"
 /protein_id="AAC66214.1"
 /db_xref="GI:2690111"
 /translation="MPYLVNNSNVOKETILKALKFFODRKSIGNNPKIFN"
 2536. 3351
 /gene="BB106"
 2536. 3351
 /gene="BB106"
 /note="similar to GB:026562 SP:P24247 GB:M83735 PID:473817
 PID:457107 percent identity: 33.62; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=1
 /product="pfs protein (pfs)"
 /protein_id="AAC66190.1"
 /db_xref="GI:2690087"
 /translation="MKIKINILIFLSIVLNSMGNSKINTPENSOSNIIISAT
 KAEIEINKIIONKKYISIEHRRKKIAIGKLMIDHNITITATGVKINTALMTSYII
 SKVISHINAGVASGIYSDKNKFKIDGVISLETYSYEDLHFGVEIGHVEHPK
 KEKANTALIRKTSKIKINITSYNGLIITGOFIDHOFORPEPEFENAIIDMSAA
 MAQVACGKIPPIIRIGISDVNNENNDDYKRLKASSASSAKIVENLILM"
 complement(3343. 3441)
 /gene="BB107"
 complement(3343. 3441)
 /gene="BB107"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB107"
 /protein_id="AAC66215.1"
 /db_xref="GI:2690112"
 /translation="WQLKGAIIKLEDEYKIKIQLNLIKILH"
 3576. 3735
 /gene="BB108"
 3576. 3735
 /gene="BB108"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB108"
 /protein_id="AAC66196.1"
 /db_xref="GI:2690093"
 /translation="WLIISFFDYNPSASNIFIKITTPKCLAKALLPLQKQFNNLPS
 FYKNIKYSRSANICK"
 3745. 3882
 /gene="BB109"
 3745. 3882
 /gene="BB109"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB109"
 /protein_id="AAC66197.1"
 /db_xref="GI:2690094"
 /translation="WASBAKIDAESKNEVNSQKSKNSKNIEVKDTGLVKAIRKS
 L"
 3911. 4315
 /gene="BB110"
 3911. 4315
 /gene="BB110"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB110"
 /protein_id="AAC66198.1"
 /db_xref="GI:2690095"
 /translation="WGRAMTPHQQLKVNLEGLKMINLLDELKIAIDGNDNNISKY
 NDLKKVDFNNSNLSIKVYLVKNSNEDKIEAKKICIKTLMHNVEYFESVCDLKDKN

gene
 CDS

gene
 CDS

QY	575	TACATATGCGTGCAGCAACATTGCAGAAATTATTTCCGTTACATTGGCAATAGTTA	634
Db	481	TACTATGCGCTGCACAAATTTACAGAAATTTCTTCACATTTTACATTTGCTAATAGTTA	540
QY	635	TAAATAGCAACTGAAAAAATAGCCTTAATATATAGCTTTTGAGAGGTTTGACCGTATTTTGA	694
Db	541	TAAATAGCAAGCTAAAAAATAGCCTTAATATATAGCTTTTGAGAGGTTTGAACTAGTTTGA	600
QY	755	ATTATATATTTATTTAATAATGATGATGACATATTTTGATTTGATGATGCTC-ATTGTAA	813
Db	661	ATTTTA-TTTAATTTACAAAATAGCGTGATTCGATTCGATTTGCAATTTGCTCAATTTGTA	719
QY	814	TTGAAAATTAGACGCTTTTGCTTTATTTATTTAATTTTA	850
Db	720	TTAAAAGAGCTTTTACTTTGTTATTTATTTTCTCTACTA	756
SUIT 4			
LOCUS	AF085741	678 bp	DNA linear BCT 24-APR-2001
DEFINITION	Borrelia garinii membrane protein p13 (p13) gene, complete cds.		
ACCESSION	AF085741		
VERSION	AF085741.1	GI:4731124	
KEYWORDS			
SOURCE	Borrelia garinii.		
ORGANISM	Borrelia garinii.		
REFERENCE	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.		
AUTHORS	1 (bases 1 to 678)		
TITLE	Noppa, L., Ostberg, Y., Lavrinovich, M. and Bergstrom, S.		
JOURNAL	P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains		
MEDLINE	Infect. Immun. 69 (5), 3323-3334 (2001)		
PUBMED	21189251		
REFERENCE	2 (bases 1 to 678)		
AUTHORS	Noppa, L., Ostberg, Y. and Bergstrom, S.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea University, Umea S-901 87, Sweden		
FEATURES	Location/Qualifiers		
source	1..678		
	/organism="Borrelia garinii"		
	/strain="ip90"		
	/db_xref="taxon:29519"		
	/chromosome="large linear"		
	121..654		
gene	/gene="p13"		
	121..654		
CDS	/gene="p13"		
	/codon_start=1		
	/transl_table=1		
	/product="membrane protein p13"		
	/protein_id="AAD8362.1"		
	/db_xref="GI:4731125"		
	/translation="MNKFLIFLIVFCAFFSSFAODDSKSFNLAGAKEFLVETNNKDD		
	SLVPELNLFLFGVIFAGADGIDGSLIFGDAVIGILITGAYLIDKDFDNNAKRNAR		
	DFKTMKGKMLAGVIMAAVYRLTEIVLPEFFANNYNRKLKNSINIALGPEPFDIN		
	MGQASALGFGLSFKFSY"		
BASE COUNT	202 a	77 c	140 g 259 t
ORIGIN			
Query Match	62.9%;	Score 542.4;	DB 1; Length 678;
Best Local Similarity	89.6%;	Pred. No. 1.4e-71;	
Matches 606; Conservative	0;	Mismatches 66;	Indels 4; Gaps 2;
QY	98	TCTATTGGGAAACGAATTTCAGATATATCTTTGAAAAAGTTAAATTTAATAGTTTAA	157

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
Db	1	TCTATTGGTAAGCCAAATTTCCGTAACAATTTGAAAAAGTTAAATTTAAATAGTTC	60					
Qy	158	AAACCTTTTAAATTCCTATTATATGTCTATATATACCACTTTTAAATAAGAGCTTT	217					
Db	61	AAGCTTTTAAATTTCAATTAATATGCTACCATTAATACAG-TTTAAATAAGGGGCTTT	119					
Qy	218	TATGAATAAATTTTAAATTTGTTGTTGCTAGACCTTTGTGTTTCTTACGTTGCTCA	277					
Db	120	TATGAATAGTTTTTAAATTTTATTTTGGTAATCTTTTGTGCTTTTCTAGTTTGTCTCA	179					
Qy	278	AGCTGATGATTTCTAAAGACGCTTTTAAATTTGGGACGGGAGAAAACTTTAGCTTATGA	337					
Db	180	A---GATGATCTCTAAAGACATTTTAATCTGGAGCGGAGAAAAATTTTGGTTATGA	236					
Qy	338	AACTGATGAAGAAGATCCTATTGTGCCATTTTATTTGAACCTTTTAAAGGTTGGCAAT	397					
Db	237	AACTATTAAGAAGATTCCTTTGACCATTTTATTTGAACCTTTTAAAGGTTGGGAT	296					
Qy	398	AGGTCCTTTGCTCAAGGAGATATCTTGAGGGGTTTCTTATTTCTTGATTTGATGACGT	457					
Db	297	AGGTTCTTTTGTCTCAAGGAGATATCTTGAGAGTTCTCTTATTTCTTGATTTGATGCGGT	356					
Qy	458	TGCTATAGGCTTAATACTTACAGAGCTTATTTAGATATCAAGCTCTTGATTAAGAATGC	517					
Db	357	TGCTATAGGCTTAATACTTACAGAGCTTATTTGACATCAAGATTTTGAATATATATGC	416					
Qy	518	TCCAAAGCCGCTTTTAAAGTGAGCTTTGGGGTAAAGGAATGATGTTGCGACGCTGACATTAC	577					
Db	417	TAAAAAGCTGATTTTAAAGTGAGCTTTGGGGGTAAAGGAATGATGTTGCGACGCTGCTTAC	476					
Qy	578	TATGCTGTGACAAATTTGACGAATTAATTTTCCGTTTACATTTGCTTAATAGTTATAA	637					
Db	477	TATGCTGTGACAAATTTGACGAATTAATTTTCCATTTACATTTGCTTAATATTAATA	536					
Qy	638	TAGGAACCTGAAAAATAGCCCTTAATATAGCTTTTGAGAGGTTTGACCTAGTTTGATAT	697					
Db	537	CAGGAAGCTGAAAAATAGCTTAATATATAGCTTTTGAGAGGATTTTGACCTAGTTTGATAT	596					
Qy	698	TAAATAGGCCAAGCTAGACGCTTTGGGTTTGAACTATCTTTCAAAAAAGTTATTAATTT	757					
Db	597	TAAATAGGCCAAGCTAGACGCTTTGGGTTTGAACTATCTTTCAAAAAAGCTATTAATTT	656					
Qy	758	TTATTTTATTTATTA 773						
Db	657	TTATTTTATTTATTA 672						
RESULT 5								
AE000789/c		27323 bp	DNA	linear	ECT 17-DEC-1997			
LOCUS	AE000789							
DEFINITION	Borrelia burgdorferi plasmid lp28-4, complete plasmid sequence.							
ACCESSION	AE000789							
VERSION	AE000789.1	GI:2690079						
KEYWORDS								
SOURCE								
ORGANISM	Borrelia burgdorferi.							
	Borrelia burgdorferi.							
	Bacteria: Spirochaetales; Spirochaetaceae; Borrelia; Borrelia							
	burgdorferi group.							
REFERENCE	1 (bases 1 to 27323)							
AUTHORS	Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.,							
	Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,							
	Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D.,							
	Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S.,							
	Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J.,							
	Venter, J.C. et al.							
TITLE	Genomic sequence of a Lyme disease spirochaete, Borrelia							
	burgdorferi							
JOURNAL	Nature 390 (6660), 580-586 (1997)							
MEDLINE	98065943							
PUBMED	9403685							
REFERENCE	2 (bases 1 to 27323)							
AUTHORS	Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A.,							
	Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,							

Matches 740; Conservative 0; Mismatches 106; Indels 5; Gaps 3;

```

OY 4 TTTTCATGSGATCCAGAAATTTAGAAATTTGCACAAAATTAAGACATATATAAAGAT 63
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6646 TTTTCATGSAATCCAGAAATTTGAAATTTTCACAAATGAAGACATGTTAAAGAT 6587
OY 64 TGAATGCTAATTTATGTCAAATCAAGAGCTATTTGGGAAGCAATTTCAAGTAA 123
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6586 TGAATGATATAATTTATGTCAAATCAAGAGCTATTTGGGAAGCAATTTCAAGCA 6527
OY 124 TACTTTGAAAAAGTTAAATTTAAATAGTTTAAAAACCTTTTAAATTCATATAT 183
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6526 TAAATTTAAAAAGTTAAATTTAAATTAATTAACCTTTTAAATTCATATATAT 6467
OY 184 GTTCTATATATACAGTTTATATAAGGTTTATGAATAAATTTATTTGTTT 243
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6466 GCTACCATATACAGTTTATATAAGGTTTATGAATAAATTTATTTGTTT 6407
OY 244 TGCTAGCCTTTGTTGTTTCTGCTTCTCAAGCTGATGATCTTAA--GCCTT 300
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6406 TGCAACCTTTTGTGTTTCTGCTTCTCAAGCTGATGATCTTAAATGTTGCT 6347
OY 301 TTAATTTGGAGCGGAGAAAACTTTTACCTATGAACTAGTAAGAAGATCCTATG 360
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6346 TTGGATGAGTGTGAGAAAACTTTTGTATGAAACTAGCAAGACAGATCCTATG 6287
OY 361 TGCAATTTTATGACCTTTTATAGGTTTGAATAGTCTTTTCTCAAGAGATA 420
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6286 TACCAATTTTATGAACTTTTATAGGTTTGAATAGTCTTCTCAAGAGATA 6227
OY 421 TTTCTGGGGGTTTCTTATTTCTGGAATTTGATGAGTGTATAGGTTTATCTTAC 480
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6226 TTTCTGGAGGTTCTCTATTTCTGATTTGATGCGGTTGATAGGCTTATCTTGGG 6167
OY 481 GAGCTTATTTGATATCAAGCTCTGTATGAATGCTCCAAAAGCGCTTTAAGTGA 540
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6166 GGGCTTATTTGGAATCAAAAGCGTGTATGATTAACAAAAGCTTTCAATGGA 6107
OY 541 CTTGGGGAAGGAGATGATTTGGCAGGTGAGTACTATGCTGATGCAAGATTTGAC 600
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6106 CTTGGGGAAGGAGATGATTTGACAGGTGATGCTGATGCTGATGCAAGATTTA 6047
OY 601 AAATTTATTTCCGTTTACATTTGCTAATAGTATTAATAGGAAACGTAAGACCT 660
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6046 AAATTTATTTCCATTTACATTTGCTAATAGTATTAATAGGAAACGTAAGACCT 5987
OY 661 AATATGCTTTGGAGGTTTGAAGCTTATGATTAATTAATGAGGCAACGCTGCTC 720
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5986 ATGATGCTTTGGAGGATTTGAACCTATTTGATGCAATGAGCAATCCAGTCTC 5927
OY 721 TTGGGTTGAACTATCTTCAAAAAAGTATTAATTTTATTTTATTAATTAATGAGTG 780
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5926 TTGGGTTGAACTATCTTCAAAAAAGTATTAATTTTATTTTATTAATTAATGAGTG 5868
OY 781 ATAGCAATTTGATTTGATGCTC-ATGTAATTTGAATAATGAGCTTTTGTATTA 839
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5867 ATGCAATTTGATTTGATGCTC-ATGTAATTTGAATAATGAGCTTTTGTATTA 5808
OY 840 TTTATATTTA 850
    || | | | | |
Db 5807 TTTCTTACTA 5797

```

RESULT 3
AF085739

LOCUS 781 bp DNA linear BCT 24-Apr-2001
DEFINITION Borrelia burgdorferi membrane protein p13 (p13) gene, complete cds.

ACCESSION AF085739
VERSION AF085739.1 GI:4731120

KEYWORDS
SOURCE

ORGANISM
Borrelia burgdorferi.
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia

REFERENCE
1 (bases 1 to 781)
Noppa, L., Ostberg, Y., Lavrinovicha, M. and Bergstrom, S.
p13, an integral membrane protein of Borrelia burgdorferi, is
C-terminally processed and contains surface-exposed domains
Infect. Immun. 69 (5), 3323-3334 (2001)
JOURNAL
MEDLINE
PUBMED
21189251
11292755
REFERENCE
2 (bases 1 to 781)
Noppa, L., Ostberg, Y. and Bergstrom, S.
Direct Submission
Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea
University, Umea S-901 87, Sweden
FEATURES
source
1. 781
/organism="Borrelia burgdorferi"
/strain="B31"
/db_xref="taxon:139"
/gene="p13"
/chromosome="large linear"
122..661
/gene="p13"
122..661
/gene="p13"
/note="similar to Borrelia burgdorferi predicted coding
region B90034 encoded by Genbank Accession Number
AB01117"

CDS

gene

BASE COUNT 223 a 98 c 153 g 307 t
ORIGIN

Query Match 65.2% Score 561.8; DB 1; Length 781;
Best Local Similarity 86.5%; Pred. No. 1.8e-74;
Matches 655; Conservative 0; Mismatches 97; Indels 5; Gaps 3;

```

OY 98 TCTATTGGGAAGCAATTTCAAGTAATCTTGAAGAAAGTTAAATTAATGTTTAA 157
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TCTATTGGGAAGCAATTTCAAGTAATCTTGAAGAAAGTTAAATTAATGTTTAA 60
OY 158 AAACCTTTTAAATTTCAATTAATGTTTACTATTAATACGATTTTAAAGGTTT 217
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 AAACCTTTTAAATTTCAATTAATGTTTACTATTAATACGATTTTAAAGGTTT 120
OY 218 TATGAATTAATTTTAAATGTTTGTGTTTGTGCTAGCCCTTTTGTGTTTCTTACCTTGC 277
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 TATGAATTAATTTTAAATTTTGTGTTTGTGCTAGCCCTTTTGTGTTTCTTACCTTGC 180
OY 278 AGCTGATGATTTCTAATA--GCCTTTTAAATTTGGAAGCGGAGAAAAAATTTAGCTTA 334
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AGCTAATGATTTCTAATAATGTTGCTTTGGGATGAGTCTGGAAGAAAAAATTTGTTTA 240
OY 335 TGAACCTAGTAAGAAAGATCTATTTGCTCCATTTTATTTGAAACCTTTTATAGGTTTG 394
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 241 TGAACCTAGTAAGAAAGATCTATTTGCTCCATTTTATTTGAAACCTTTTATAGGTTTG 300
OY 395 AATAGGTTCTTTGCTCAAGAGATATCTTGGGGGTTTCTTATTTCTGATTTGATGC 454
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 AATAGGTTCTTTGCTCAAGAGATATCTTGGAGGTTTCTTATTTCTGATTTGATGC 360
OY 455 AGTTGATTAAGGTTAATTAATCTTACAGAGCTTATTTAGATATCAAGCTCTGATTAAGAA 514
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 GATTGATTAAGGCTTATTAATCTTACAGAGCTTATTTAGATATCAAGAGCTCTGATTAAG 420
OY 515 TGCTCAAAAAGCCCTTTTAACTGAGCTTGGGCTGAAGGATGATGTTGCGAGGTGAGT 574
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 TACTAAAAAAGCTCTTTTCAATGAGCTTGGGCTGAAGGAGTTATGTTAGCAGGTGTGCT 480

```

```

putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0028"
/protein_id="AAC6428.1"
/db_xref="GI:2687919"
/translation="MKQRYENFKRRLILNLLPILLACSSSEIFSOGLNOKIKHEY
NILGSSSPRGLISVETLYIAMHLFKKENGKIEKIDLSNSYEFLINDIYNISGKTYL
AONKEEELVECELNGDKMTLKKRPLKAYKPLKSGRGVKEAYILAIADKNKREKIFD
LOGSDKTPPOATENGDKRQIOISNEELTNGSLKIQMNNNTYTNIDYQOAKETPILK
TSIRGSSSELYVTGCGYNNIDTRFKVYSTNNNTTTFITODEVGEFSYFAREFNAIL
IGSNGFAEFKNEGIFALRAPSVEPGAVNGSQLSKTGLINDIIPVSNNTIYILTO
GGLMKLENKRLKE"
/complement(1776. .2207)
/gene="BB0029"
/complement(1776. .2207)
/gene="BB0029"
/note="similar to SP:P17868 GB:X03897 PID:40141
PID:1303831 GB:AL009126 percent identity: 32.58;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC6423.1"
/db_xref="GI:2687914"
/translation="MLSKIFVDADSCNFIIKLOKFIHLNKSRLIIVSNKPFSEIT
ENVALEVVDNSFILLADRSIVITRDIPFAKLILDGVKVMNDEGRIPDINNIN
LYFRSKINPILKIKIKKYDEDFNKSRYEKETTNYSLFPS"
/complement(2197. .2832)
/gene="BB0030"
/complement(2197. .2832)
/note="similar to GB:Z36269 PID:1049113 percent identity:
30.95; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="signal peptidase I (lepB-1)"
/protein_id="AAC6422.1"
/db_xref="GI:2687913"
/translation="MAPYLTFEORLLRRKORVPEKVFLETLINFEPTKPYLQIFMI
KSNEMLPITTKNASLFFVATHTTSFFIPLKANDIYLFEDFRLSNFLTLTKDFEFLN
KIFKASIKVSRILAVOGDSYVRGLVNLVNNKIDNFYLNGLNLSYKRLKDNFEFTDL
VICKETLKNEFFLNDMLSVLNDSRIFGPINKNAIVSEFLVKVDYKIVK"
/complement(2834. .3814)
/gene="BB0031"
/complement(2834. .3814)
/note="similar to PID:1652260 percent identity: 35.98;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="signal peptidase I (lepB-2)"
/protein_id="AAC6421.1"
/db_xref="GI:2687912"
/translation="MYLRLDKFASFLVHVEKYLTYRRKKYFCKLRBAKRGFLNLE
LFDEVAALFVLIYQVQVQAVKIPSGSMENYLOGLDFVDRKSGYGPDLGLPKFN
GAPAESEDIILFENPEYKSGVDFIDIRLLIYMTLSFIDLRDEGPNVRYLVR
GAPADGKYVRNNGKAYIKRGEENFILEDSDWDVDNFKTKIYANEDGIGTYDFA
MVALSOLNINSSTPDFSFDVRMIDKEFERLEKYLASMPYVDVYMEALMRYD
GIYPDGIILPGLDRNDSHGRFFGVINKKVLGRLLIYLFPSRVGFI"
/3885. .5357
/gene="BB0032"
/3885. .5357
/gene="BB0032"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0032"
/protein_id="AAC6427.1"
/db_xref="GI:2687918"
/translation="MKVEDNSNVELFEVKLPLIGIEPKVYVFLTLLILSLISTL
IILPKFNPGAYLKINSNIENYIYLNEKYIGRPLLNKYINATCEVLAKAKRGFETYE

```

```

ORIKHNKFCNGYSLOINLELVDEPEIKIKOROKELSIWIKIKINENTRLIIPVSELS
SELKHPYIKKFKELDSIPYLNSTEMPEFNSYAIYSIODNSNOEINSLKTFN
DLNRAIFPELENDLKDILKTRNAPWKTAKTDIANDIOLISNENIKTLPGFKF
INSNKIEKIÖNELNSLDSKINSKISTYINAKFEILOEONVTKYIEÖDLPKEPKALN
KNEMLIKOEÖLVENLKNENÖIGLEAATIGISYFSAITANNYSKRLPTGFEARLPIIS
ÖEMELYOKEPNKNPLNINISKRVGMFMOMSSNEIAIFENENKPYSENSNFSYLI
TEIRFYSOÖNNLINASTRKASFLKMSSPNIGFRILVSE"
/5369. .5821
/gene="BB0033"
/5369. .5821
/gene="BB0033"
/note="similar to GP:2073503 percent identity: 47.68;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="small protein (smpB)"
/protein_id="AAC6420.1"
/db_xref="GI:2687911"
/translation="MNTNNLLLENKAKFNFYIEKISGIVLKGTEVKSIRAKLSTF
NSFASIKKEELMLENLHVSRYKEGNFNHDELRRKLILKRELÖRLKFKKEKGYT
LIPISFYLKSITKYEVGICGGKKLYDKRELKOKSIKKDLSREIKYK"
/complement(5892. .6431)
/gene="BB0034"
/complement(5892. .6431)
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BB0034"
/protein_id="AAC6426.1"
/db_xref="GI:2687917"
/translation="MNKLLIFLAFCVFPSSPAONDSKNGAFSGAGEKILYVETSK
QDPYPLNLNPLFGISGSRFQOGDILGSSLITIGFPAVIGILNAGYIADIRALDSITK
KAAFWTWGKGVLAGVYVMAVTRULTEIILPTFANSTRKLKNSLNAVLAGFEPSPD
VAMGSSALGFEELSPKSY"
/complement(6499. .8379)
/gene="BB0035"
/complement(6499. .8379)
/note="similar to GP:1405462 percent identity: 31.66;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA topoisomerase IV (parC)"
/protein_id="AAC6419.1"
/db_xref="GI:2687910"
/translation="MDIRTVLAKDNFLQSSYVIKDRAISAVDGFKPYORRIHSLFE
MHGDNFHVAVVNGTMRKYHPHGDSIYEALVNIANKDLFTEOGFNGNLFTGPASA
SRYTECRLLPLAFDVLYSKEITTYESSYDGRNNEPLYPAKIPVTLIOGSGIAGMA
AKIILPHNNEILNAYKSLLEGSDYIYDPFGCGIVDNEYADNGKYLVAKIEFTID
EKTIVIRELPREYTESLISSTERKIRKNYIKVSSINDPTAEVAIEISLPGVYASE
VLEKLYHTNQÖISISVNLLESEKYPVAVYTKDLIRKHAHLOILMELEÖSKI
LEKIEFKTLEÖQIFIEKIKYKLEETISKEENILSISLEVLHKSFSREVAKEDVENL
LKIPRLKISLEÖIDNSKÖIKLANKLEKINSINSISIGSYINFDLILAKESHOR
KTKISLKSINWKEIATANMVMYVLAGPAGTSLFOEGEFGMSYVYKILIVFRENYS
VLKNIEDFTPIDKRVQALYVIDNNSKQÖISITIFPNLDPYIYKRRKIDKPTIDKY
VLEFGEÖNEFVDESINPEVEFSTNKDILVKRIEDNEVKSRSISIGKRISSNLLKVK
FK"
/complement(8379. .10178)
/gene="BB0036"
/complement(8379. .10178)
/gene="BB0036"
/note="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018
PID:467396 percent identity: 37.98; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA topoisomerase IV (parE)"
/protein_id="AAC6418.1"

```

Query Match 74.4%; Score 641.4; DB 1; Length 14102;
Best Local Similarity 87.0%; Pred. No. 1.1e-86;

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:45:32 ; Search time 1923.68 Seconds
(without alignments)
13040.962 Million cell updates/sec

Title: US-09-508-487-20

Perfect score: 862
Sequence: 1 gatttcattgcacccag.....atatattctctcgtctaa 862

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl1:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_pl:*
25: em_pi:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_higo_hum:*
40: em_higo_mus:*
41: em_higo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	764	88.6	777	1	AF085740	AF085740 Borrelia
2	641.4	74.4	14102	1	AE001117	AE001117 Borrelia
3	561.8	65.2	781	1	AF085739	AF085739 Borrelia
4	542.4	62.9	678	1	AF085741	AF085741 Borrelia
5	100	11.6	27323	1	AE000789	AE000789 Borrelia
6	94.4	11.0	29766	1	AE000786	AE000786 Borrelia
7	88.2	10.2	52971	1	AE001584	AE001584 Borrelia
8	71.4	8.3	43993	2	AC116965	AC116965 Dictyoste
9	67.4	7.8	14950	6	AX346159	AX346159 Sequence
10	67.2	7.8	349980	6	AX344568	AX344568 Sequence
11	67	7.8	5930	6	AX345419	AX345419 Sequence
12	67	7.8	6012	6	AX251052	AX251052 Sequence
13	67	7.8	6012	6	AX344183	AX344183 Sequence
14	67	7.8	6012	6	AX348580	AX348580 Sequence
15	66.4	7.7	129360	2	AC117079	AC117079 Dictyoste
16	65.8	7.6	9742	6	AX348912	AX348912 Sequence
17	65.4	7.6	115489	2	AC117072	AC117072 Dictyoste
18	65.2	7.6	11036	6	AX323649	AX323649 Sequence
19	65.2	7.6	11036	6	AX323649	AX323649 Sequence
20	64.8	7.5	104992	2	AC005504	AC005504 Plasmodu
21	64.6	7.5	131271	2	AC015927	AC015927 Homo sap
22	64.4	7.5	6590	6	AX251927	AX251927 Sequence
23	64.4	7.5	6590	6	AX346285	AX346285 Sequence
24	64.4	7.5	6590	6	AX348730	AX348730 Sequence
25	64.4	7.5	19087	6	AX345695	AX345695 Sequence
26	63.8	7.4	17967	6	AX458472	AX458472 Sequence
27	63.6	7.4	349980	6	AX345917	AX345917 Sequence
28	63.4	7.4	5815	6	AX344565	AX344565 Sequence
29	63.4	7.4	16766	6	AX346137	AX346137 Sequence
30	63.4	7.4	349980	6	AX347059	AX347059 Sequence
31	63.4	7.4	349980	6	AX344551	AX344551 Sequence
32	63.4	7.4	349980	6	AX344552	AX344552 Sequence
33	63.2	7.3	7441	6	AX251450	AX251450 Sequence
34	63.2	7.3	7441	6	AX348445	AX348445 Sequence
35	63.2	7.3	7479	6	AX339173	AX339173 Sequence
36	63.2	7.3	13784	6	AX348449	AX348449 Sequence
37	63.2	7.3	19380	6	AX252128	AX252128 Sequence
38	63	7.3	106434	3	AC117080	AC117080 Dictyoste
39	63	7.3	163290	2	AL845514	AL845514 Danio rer
40	62.8	7.3	133501	2	AC116956	AC116956 Dictyoste
41	62.6	7.3	6716	6	AX346684	AX346684 Sequence
42	62.4	7.2	5963	6	AX281502	AX281502 Sequence
43	62	7.2	6509	6	AX251781	AX251781 Sequence
44	62	7.2	6509	6	AX344185	AX344185 Sequence
45	61.8	7.2	5884	6	AX347067	AX347067 Sequence

ALIGNMENTS

```

RESULT 1
AF085740          AF085740          777 bp      DNA      linear      BCT 24-APR-2001
LOCUS            Borrelia afzelii membrae protein p13 (p13) gene, complete cds.
DEFINITION
ACCESSION        AF085740
VERSION          AF085740.1 GI:4731122
KEYWORDS
SOURCE
ORGANISM         Borrelia afzelii.
                  Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
                  burgdorferi group.
REFERENCE
AUTHORS          Noppa,L., Ostberg,Y., Lavrinovitcha,M. and Bergstrom,S.
TITLE            p13, an integral membrane protein of Borrelia burgdorferi, is
                  C-terminally processed and contains surface-exposed domains

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:45:32 ; Search time 1693.82 seconds
(without alignments)
13040.962 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759
Sequence: 1 atgttaagaatgaat.....atctgtatgaatgggtg 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_bg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sy:*
- 12: gb_un:*
- 13: gb_un:*
- 14: gb_vt:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vt:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rtd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	753.2	99.2	14102	1	AE001117	AE001117 Borrelia
2	703.8	92.9	781	1	AF085739	AF085739 Borrelia
3	550.8	72.6	777	1	AF085740	AF085740 Borrelia
4	513.4	67.6	678	1	AF085741	AF085741 Borrelia
5	104.8	13.8	52971	1	AE001584	AE001584 Borrelia
6	92.6	12.2	29765	1	AE000786	AE000786 Borrelia
7	90.6	11.9	27323	1	AE000789	AE000789 Borrelia
C 8	67.8	8.9	173540	2	AC119687	AC119687 Rattus no
C 9	66	8.7	28601	1	AE000784	AE000784 Borrelia
C 10	62	8.2	236130	9	AL592044	AL592044 Human DNA
11	57	7.5	106434	3	AC117080	AC117080 Dictyoste
12	56.6	7.5	6852	6	AX251949	AX251949 Sequence
13	56.6	7.5	6852	6	AX348744	AX348744 Sequence
C 14	56.4	7.4	193379	2	AC121387	AC121387 Rattus no
C 15	56	7.4	133501	2	AC116956	AC116956 Dictyoste
16	55.6	7.3	18855	6	AX345512	AX345512 Sequence
17	55	7.2	53561	1	AE000790	AE000790 Borrelia
18	54.6	7.2	7106	6	AX346380	AX346380 Sequence
19	54.2	7.1	73334	6	AX347027	AX347027 Sequence
20	54.2	7.1	73334	6	AX356494	AX356494 Sequence
21	54	7.1	18624	6	AX346604	AX346604 Sequence
22	54	7.1	60604	2	AC023466	AC023466 Homo sapi
23	54	7.1	73377	2	AC117077	AC117077 Dictyoste
24	53.8	7.1	396	1	BORS3G	LA1151 Borrelia bu
C 25	53.8	7.1	115489	2	AC117072	AC117072 Dictyoste
C 26	53.8	7.1	124820	2	AC117073	AC117073 Dictyoste
C 27	53.6	7.1	33068	3	AC115682	AC115682 Dictyoste
C 28	53.4	7.0	2532	1	CJ009019	U09019 Campylobact
29	53.4	7.0	30726	2	AC117269	AC117269 Dictyoste
30	53.4	7.0	155932	1	CJ11168X6	AL139079 Campyloba
31	53.4	7.0	349980	6	AX344559	AX344559 Sequence
32	53.2	7.0	156533	2	AC117070	AC117070 Dictyoste
33	53.2	7.0	163034	2	AC099432	AC099432 Rattus no
34	53.2	7.0	349980	6	AX344566	AX344566 Sequence
C 35	53	7.0	3606	3	DD11093A	M19469 Dictyosteli
C 36	52.8	7.0	7921	3	AF153362	AF153362 Dictyoste
37	52.6	6.9	12025	6	AX346200	AX346200 Sequence
38	52.6	6.9	84805	2	AC116918	AC116918 Dictyoste
39	52.6	6.9	164399	3	PFMAL3P6	Z98551 Plasmodium
40	52.2	6.9	6012	6	AX251975	AX251975 Sequence
41	52.2	6.9	6012	6	AX344367	AX344367 Sequence
42	52.2	6.9	6012	6	AX348760	AX348760 Sequence
43	52.2	6.9	12507	6	AX345201	AX345201 Sequence
44	52	6.9	6963	6	AX345881	AX345881 Sequence
45	51.8	6.8	5586	6	AX348390	AX348390 Sequence

ALIGNMENTS

RESULT 1	AE001117/c	14102 bp	DNA	linear	BCT 15-DEC-1997
LOCUS	AE001117				
DEFINITION	Borrelia burgdorferi	(section 3 of 70)	of the complete genome.		
ACCESSION	AE001117				
VERSION	AE001117.1	GI:2687907			
KEYWORDS					
SOURCE					
ORGANISM	Borrelia burgdorferi.				
	Borrelia burgdorferi				
	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia				
	Burgdorferi group.				
	1 (bases 1 to 14102)				
REFERENCE	Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.,				
AUTHORS	Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K.,				
	Gwinn,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D.,				

FEATURES		
TITLE	Journal	
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA	
REFERENCE	Submitted (12-DEC-1997) The Institute for Genomic Research, 9712	
AUTHORS	Location/Qualifiers	
	1. .14102	
	/organism="Borrelia burgdorferi"	
	/db_xref="taxon:139"	
	95. .733	
	/gene="BB0027"	
	95. .733	
	/note="hypothetical protein; identified by Glimmer; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="B. burgdorferi predicted coding region BB0027"	
	/protein_id="AAC66429.1"	
	/db_xref="GI:2687920"	
	/translation="MKRIYRITLLAVLLIGVNIKKIAAANIDRHTSTGIDLSVCI PIFYNDLSKAPYTNLYPGGIGAIKTYOHILNLAIGLELMFEPDINHSFNLPDS SVGKIFYSPIITSINIFIDIGELFQIPVTNIGFSLNTYGDNRNNTITNRTDALPT ISFGSGILMNFNKNMAGATSWMMFEFGNSAKMAHFALVSLSVYVNVKL"	
	741. .1790	
	/gene="BB0028"	
	741. .1790	
	/note="hypothetical protein; identified by Glimmer; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="B. burgdorferi predicted coding region BB0028"	
	/protein_id="AAC66428.1"	
	/db_xref="GI:2687919"	
	/translation="MKQYENYFKKRIILNLLIFLLACSSSEIFISQGLNLIKIKHE NILGSSSPGICISLIGETLYIAAMHLPKENGKTEKIDLSNYPEINDIVINSGTYL AONKEEELGECINGKDWTLKPKKPLKAYKELSGVDGYKEAVIILADNNRKEIPD LOGSDKTPPOATENDKRYQISNEENLITGSLSKIMWNNNTYITIDYQAKERPIIK TISIRGSEVLVMTGGIINNLDPKERKVSNTNNYTPITIDDEVGEFSSTFAEEEDAL IGSNNGFAEFTKKEGIFALRAPSVEPGAVNGSOLSKGLANDIIPVSNNTYIILQ GKGLMKLENKRIKE"	
	complement(1776. .2207)	
	/gene="BB0029"	
	complement(1776. .2207)	
	/note="similar to SP:P17868 GB:X03897 PID:40141 PID:1303831 GB:AL009126 percent identity: 32.58; identified by sequence similarity; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="conserved hypothetical protein"	
	/protein_id="AAC66423.1"	
	/db_xref="GI:2687914"	
	/translation="MLSKITVDADSCNFKIIPLOKFIILNHSKLIYVSNKFFSELT ENVALEVVANDVDFILNLADYSLVITRIDIFFAKLIDLGVKVNDRGRIFDINNIN LYFRSKINFLKIKIRKYDEDFNKSRYEKRTITNFYSLFS"	
	complement(2197. .2832)	
	/gene="BB0030"	
	complement(2197. .2832)	
	/note="similar to GB:Z36269 PID:1049113 percent identity: 30.95; identified by sequence similarity; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="signal peptidase I (lepB-1)"	
	/protein_id="AAC66422.1"	
	/db_xref="GI:2687913"	
	/translation="MAPYLFEQRLRKORRYEFTFLINFEFTKVLQIEMI KSNEMPTITKNSLFEVAHITLSEFPIKMDVIYEDPELSNFFLLIKOFEFLN KIKRASKYSKRAAVQGSYVRGLNVLNKKDTPFYLNGVLSYVTLNDFENDE VICFTLKNEFFLLDNLSVLDSRTIFGFINNALVSFLVKVVDIKYK"	
	complement(2834. .3814)	
	/gene="BB0031"	
	complement(2834. .3814)	
	/note="similar to PID:1652260 percent identity: 35.98; identified by sequence similarity; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="signal peptidase I (lepB-2)"	
	/protein_id="AAC66421.1"	
	/db_xref="GI:2687912"	
	/translation="MYIRLDRKFAFLVHVEYLYYRKKRKYFCKLRKRRGFLNF LFDVAALFVAVINQYEVQYKIPSGSMENITLQIDFLFVDKFSYGPDLPELFIN GKRAPESDIIIFENPEYKSGVFPDIFORILYMLTSLFDLDRDYGNNPVLVR GARADGKIVRNNKAVIKREGSENFLEDSYDLVDKNEKIKKIYANEDYGIYGPFA MFVALSOLNINLSSTPDFSFDFRMDIDREFEELKYSAPFVYDYVYKAIMRBY GIYVPGCYITLPIGDNRDNSHDGKFFGVYKNNKVLGRLIITLFPSSVGF"	
	3885. .5357	
	/gene="BB0032"	
	3885. .5357	
	/note="hypothetical protein; identified by Glimmer; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="B. burgdorferi predicted coding region BB0032"	
	/protein_id="AAC66427.1"	
	/db_xref="GI:2687918"	
	/translation="MKQYENYFVKRIKPIIGTEPVYVFIITLILSLISLT IILPKFNGGAVLAKINSINENTYIYLEKTIQRTPLKRYINATEGVILAKRMRKPYE QRIKIHKKFPGNYSLOINLELVDEPKIKOROLESLIMVAKININENTKLIIPVSLIS SELKEHPKYIKKFLKDSIPYLNSTEMKIDLSYKAYSIDQNSQDELTWNLKFN DLNRAIFMFLNLDKDLITKNAFVKTLLAKTIDENQOLSKNEKINIKLPGFK INSNKTEKIONELNSDKNSKINISKYNKEPLIOQNTYKKEODFLKENPKMALN NKENLKEQLYVDENYLNKNOICLNPATITGTSYSPATEVNNYSKRLPGFKARPLIS QEWELVKEPKNKNPLNINETSRYGVGNLQNSSPENIATFKNEKNFYSENSNYSLI TEIRTSQOQNNNLNASTKASFLKMWSSPNIIGRLIVSKP"	
	5369. .5821	
	/gene="BB0033"	
	5369. .5821	
	/note="similar to GP:2073503 percent identity: 47.68; identified by sequence similarity; putative"	
	/codon_start=-1	
	/transl_table=11	
	/product="small protein (smpB)"	
	/protein_id="AAC66420.1"	
	/db_xref="GI:2687911"	
	/translation="MTNNLLLENKRAKAFNYFEKISGIVLKGTEVKSIAKKLSF NNSPASIKKEELMELNHSVKEGNFENDELRLLKRELQDLKFKREKEGYT LIPISFLKSIITVEVGICGKKLYKRELKAKSKIKKOLSRIRIKYK"	
	complement(5892. .6431)	
	/gene="BB0034"	
	complement(5892. .6431)	
	/note="hypothetical protein; identified by Glimmer; putative"	
	/codon_start=-1	
	/transl_table=11	


```
/product="B. burgdorferi predicted coding region BB0034"
/protein_id="AAC66426.1"
/db_xref="GI:2687917"
/translation="MNKLIFVLATFCVSSFAQANDSKNGAFSGAEKLLVETSK
QDPVLPFLNLFLGIGSFAOGDILIGFPAVGIGLILAGAYDIALDGIITK
KAAPFMTWKGCVMLAGVVTMAVTRLEIILPFTFANSYNRKLKNSLNLALGFEPSFD
VAMGSSALGFEISPKSY"
complement(6499..8379)
/gene="BB0035"
complement(6499..8379)
/gene="BB0035"
complement(8379)
/gene="BB0035"
/feature="similar to GP:1405462 percent identity: 31.66;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="DNA topoisomerase IV (parc)"
/protein_id="AAC66419.1"
/db_xref="GI:2687910"
/translation="MDIRVLKDNFLQSYVYIKDRATASVVDGKPYORRIISLFE
MHDGNFHVAVNGVMTKYPHGDSTYEALVNTAKPLFIEKQGNFNLGDPASA
SRVIECRPLAFVLYSKETITHESSIDGRNEDLPKIPVLILGSEIAGVMA
AKILPHNNEILNVAKSELDESVDIYDPFGIVDNEADGKXVLRKLETTID
EKIVIRLPGETESLISIEKAIKRNKIVKSSINDPENAIELESLRGVASE
VIEKLYHTNCOISISVNLISERYPVVYIKDLIKFHAHLKLEKMELELOSKI
LEKIFVKTLEOPIETKLYKLETSRENNILSTLSVLRHKSSESPREKEDENL
LKIPRKISLEFDIDKNSDKILINKELKSINSNISIRGYSINFDILAKTSKHOK
KTKISLISKVNEKELATNMKRVYLNLAGFAGTSLDEFGNAYSDKILVERNSY
VLKNIEDFTFDKKNVCAVLDINNKSQIISILYFNRLDEYVYKRIKIDKFIIDKY
YEFGENDEFVDFSLNPEFEFSTNKDVKRIEIDNFWAKRSISGKRISNNLEKVK
FK"
complement(8379..10178)
/gene="BB0036"
complement(8379..10178)
/gene="BB0036"
/feature="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018
PID:467396 percent identity: 37.98; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="DNA topoisomerase IV (parc)"
/protein_id="AAC66418.1"
Query Match 99.2% Score 753.2; DB 1; Length 14102;
Best Local Similarity 99.6% Pred. No.2,1e-117;
Matches 755; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 421 GCTTATACCTGGGGGCTTATTTGGATATCAAGCGCTTGATGATTAATAAAGC 480
|||||
Db 6180 GCTTATACCTGGGGGCTTATTTGGATATCAAGCGCTTGATGATTAATAAAGC 6121
QY 481 TCGTTTCAATGACTTGGGTAAGGACTTATGTAGCAGGTGGTACTATGCTGT 540
|||||
Db 6120 TCGTTTCAATGACTTGGGTAAGGACTTATGTAGCAGGTGGTACTATGCTGT 6061
QY 541 GACAGATTACAGAAATATCTTCCATTATACATTTGCTAATAGTATATAGAAAGCT 600
|||||
Db 6060 GACAGATTACAGAAATATCTTCCATTATACATTTGCTAATAGTATATAGAAAGCT 6001
QY 601 AAAAAATACCTTAATAGTACCTTATAGAGATTGAACTAGTTTGATGCTGATGGG 660
|||||
Db 6000 AAAAAATACCTTAATAGTACCTTATAGAGATTGAACTAGTTTGATGCTGATGGG 5941
QY 661 CCAATCCAGTGCCTCTGGGTTGAACGTGCTTCAAAAAAGCTATTAATTTATTTATT 720
|||||
Db 5940 CCAATCCAGTGCCTCTGGGTTGAACGTGCTTCAAAAAAGCTATTAATTTATTTATT 5881
QY 721 AAAAAATGGGTGATTGCAATTCGTATTTGAATGGCT 758
|||||
Db 5880 AAAAAATGGGTGATTGCAATTCGTATTTGAATGGCT 5843

RESULT 2
AF085739 781 bp DNA linear BCT 24-APR-2001
LOCUS AF085739
DEFINITION Borrelia burgdorferi membrane protein p13 (p13) gene, complete cds.
ACCESSION AF085739
VERSION AF085739.1 GI:4731120
KEYWORDS
SOURCE
ORGANISM Borrelia burgdorferi.
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
burgdorferi group.
1 (bases 1 to 781)
Noppa, L., Ostberg, Y., Lavinovitch, M. and Bergstrom, S.
p13, an integral membrane protein of Borrelia burgdorferi, is
C-terminally processed and contains surface-exposed domains
infect. Immun. 69 (5), 3323-3334 (2001)
JOURNAL
MEDLINE
PUBMED
21189251
11292755
2 (bases 1 to 781)
Noppa, L., Ostberg, Y. and Bergstrom, S.
Direct Submission
Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea
University, Umea S-901 87, Sweden
FEATURES
source
1..781
/organism="Borrelia burgdorferi"
/db_xref="B31"
/strain="B31"
/chromosome="large linear"
/gene="p13"
122..661
/gene="p13"
122..661
/feature="similar to Borrelia burgdorferi predicted coding
region BB0034 encoded by Genbank Accession Number
AE001117"
/codon_start=1
/transl_table=11
/product="membrane protein p13"
/protein_id="ABD28360.1"
/db_xref="GI:4731121"
/translation="MNKLIFVLATFCVSSFAQANDSKNGAFSGAEKLLVETSK
QDPVLPFLNLFLGIGSFAOGDILIGFPAVGIGLILAGAYDIALDGIITK
KAAPFMTWKGCVMLAGVVTMAVTRLEIILPFTFANSYNRKLKNSLNLALGFEPSFD
VAMGSSALGFEISPKSY"
BASE COUNT 223 a 98 c 153 g 307 t
ORIGIN
```

Query Match 92.9%; Score 705.2; DB 1; Length 781;
 Best Local Similarity 99.6%; Pred. No. 5.1e-109;
 Matches 707; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 49 TCTTTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTTAACTTTAA 108
 |||||||
 Db 1 TCTTTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTTAACTTTAA 60

QY 109 AAACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 168
 |||||||
 Db 61 AAACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 120

QY 169 TATGAATPAACCTTTTAAATTTTGTGTTGGCAACCTTTTGTGTTTCTAGCTTTGCTCA 228
 |||||||
 Db 121 TATGAATPAACCTTTTAAATTTTGTGTTGGCAACCTTTTGTGTTTCTAGCTTTGCTCA 180

QY 229 AGCTAATGATCTTAAAAATGCTGGCTTTGGGATGATGCTGAGAAAAACCTTTGGTTT 288
 |||||||
 Db 181 AGCTAATGATCTTAAAAATGCTGGCTTTGGGATGATGCTGAGAAAAACCTTTGGTTT 240

QY 289 TGAAGACTAGCAAGAGATCTATTTGACATTTTATGAACTTTTATAGGGTTTG 348
 |||||||
 Db 241 TGAAGACTAGCAAGAGATCTATTTGACATTTTATGAACTTTTATAGGGTTTG 300

QY 349 AATAGGCTCTTTGCTCAAGAGATTTCTGAGGTTCTCTTATCTTGATTTGATGC 408
 |||||||
 Db 301 AATAGGCTCTTTGCTCAAGAGATTTCTGAGGTTCTCTTATCTTGATTTGATGC 360

QY 409 GGTGGTATAGGGCTTATCTTGGGGGGCTTATTTGGATTCAAAAGGGCTGATNGTAT 468
 |||||||
 Db 361 GGTGGTATAGGGCTTATCTTGGGGGGCTTATTTGGATTCAAAAGGGCTGATNGTAT 420

QY 469 TACTAAAAAGCTCTTTTCAATGACTTGGGTAAGGAGCTTATGTAGCAGGTGTGT 528
 |||||||
 Db 421 TACTAAAAAGCTCTTTTCAATGACTTGGGTAAGGAGCTTATGTAGCAGGTGTGT 480

QY 529 TACTATGGCTGTGACAGATTAACAGAAATTAATTTCTTCATTTACATTTGCTAATAGTTA 588
 |||||||
 Db 481 TACTATGGCTGTGACAGATTAACAGAAATTAATTTCTTCATTTACATTTGCTAATAGTTA 540

QY 589 TAATAGAGAGCTAAAAATATACCTTAATGCTTATAGGAGATTTGAACTAGTTTGA 648
 |||||||
 Db 541 TAATAGAGAGCTAAAAATATACCTTAATGCTTATAGGAGATTTGAACTAGTTTGA 600

QY 649 TGTTCATATGGGCCAATCCAGTGTCTGTGGGTTTGAACGTCTTTCAAAAAAGCTATTTA 708
 |||||||
 Db 601 TGTTCATATGGGCCAATCCAGTGTCTGTGGGTTTGAACGTCTTTCAAAAAAGCTATTTA 660

QY 709 AATTTATTTATTAACAAAAATGGGTGATTTGCAATTCCTGATTGAATGGGT 758
 |||||||
 Db 661 AATTTATTTATTAACAAAAATGGGTGATTTGCAATTCCTGATTGAATGGGT 710

RESULT 3
 AF085740 777 bp DNA linear BCT 24-Apr-2001
 LOCUS Borrelia afzelii membrane protein P13 (P13) gene, complete cds.
 DEFINITION AF085740
 ACCESSION AF085740.1 GI:4731122
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Borrelia afzelii.
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.
 REFERENCES
 1 (bases 1 to 777)
 Noppa, L., Ostberg, Y., Lavrinovich, M. and Bergstrom, S.
 P13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains Infect. Immun. 69 (5), 3323-3334 (2001)
 JOURNAL MEDLINE 21189251
 PUBMED 11292755
 REFERENCE 2 (bases 1 to 777)

AUTHORS Noppa, L., Ostberg, Y. and Bergstrom, S.
 TITLE Direct Submission
 JOURNAL Submitted (24-Aug-1998) Microbiology, Medical Microbiology, Umea University, Umea S-901 87, Sweden
 FEATURES
 source
 1. 777
 /organism="Borrelia afzelii"
 /strain="ACAT"
 /db_xref="taxon:29518"
 /chromosome="large linear"
 121..657
 /gene="p13"
 121..657
 /gene="p13"
 /codon_start=1
 /transl_table=1
 /product="membrane protein P13"
 /protein_id="AAB28361.1"
 /db_xref="GI:4731123"
 /translation="MNKFLIVLAFCVSSFPQADDSKSFNLAGKELIAYETSK
 DPVPELNLFLFGIGSFAQGDILGFLILGDAVIGLILGATLIDKADKAPK
 AAFKWTWGMKMLAGAVTMAVTRLEIILIPFPANSYNRLKNSLNIAAGGEPSPDI
 NMQQASALGFEISFKRSY"
 BASE COUNT 228 a 84 c 148 g 316 t 1 others
 ORIGIN

Query Match 72.6%; Score 550.8; DB 1; Length 777;
 Best Local Similarity 87.9%; Pred. No. 4e-83;
 Matches 624; Conservative 0; Mismatches 82; Indels 4; Gaps 2;

QY 50 CTATTTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTAATTAATTA 109
 |||||||
 Db 1 CTATTTGGGAAGCAATTTCAAGCAATTAATTTGAAAAAGTTAAATTAATTAATTA 60

QY 110 AACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 169
 |||||||
 Db 61 AACCTTTTAAATTCATTAATATGCTACCATAGTACAGTTTATATAAGGGGTTT 120

QY 170 ATGAATAAATTTTAAATTTTGTGTTGGCAACCTTTGTGTTTCTAGCTTTGCTCAA 229
 |||||||
 Db 121 ATGAATAAATTTTAAATTTTGTGTTGGCAACCTTTGTGTTTCTAGCTTTGCTCAA 180

QY 230 GCTAATGATTTTAAAAATGCTGGCTTTGGGATGATGCTGAGAAAAACCTTTGGTTAT 289
 |||||||
 Db 181 GCTAATGATTTTAAAAATGCTGGCTTTGGGATGATGCTGAGAAAAACCTTTGGTTAT 237

QY 290 GAACTAGAGCAAGAGATCTATTTGACATTTTATTAAGCACTTTTATAGGTTTGA 349
 |||||||
 Db 238 GAACTAGAGCAAGAGATCTATTTGACATTTTATTAAGCACTTTTATAGGTTTGA 297

QY 350 ATAGGCTCTTTGCTCAAGAGATATCTTGGAGGTTCTTATCTTGGATTGGATGCG 409
 |||||||
 Db 298 ATAGGCTCTTTGCTCAAGAGATATCTTGGAGGTTCTTATCTTGGATTGGATGCA 357

QY 410 GTTGGTATAGGGCTTATATCTTGGGGGCTTATTTGGATATCAAGCGCTGATGAT 469
 |||||||
 Db 358 GTTGGTATAGGGCTTATATCTTGGGGGCTTATTTGGATATCAAGCGCTGATGAT 417

QY 470 ACTAAAAAGCTGCTTTTCAATGACTTGGGTAAGGAGTATTTAGCAAGCGTGTGT 529
 |||||||
 Db 418 GCTCAAAAAGCGCTTTTAACTGAGCTTGGGTAAGGAGATGTTGGCAGGTGCA 477

QY 530 ACTATGCTGTGACAGATTAACAGAAATTAATCTTCCATTTACATTTGCTAATAGTTAT 589
 |||||||
 Db 478 ACTATGCTGTGACAGATTAACAGAAATTAATCTTCCATTTACATTTGCTAATAGTTAT 537

QY 590 AATAGGAGCTAAAAAATAGCTTAATGTACCTTATAGGAGATTTGAACTAGTTTGA 649
 |||||||
 Db 538 AATAGGAGCTAAAAAATAGCTTAATGTACCTTATAGGAGATTTGAACTAGTTTGA 597

QY 650 GTTGCATGGGCCAATCCAGTGTCTTGGGTTTGAACGTCTTTCAAAAAAGCTATTTA 709
 |||||||
 Db 598 GTTGCATGGGCCAATCCAGTGTCTTGGGTTTGAACGTCTTTCAAAAAAGCTATTTA 657

gene /protein_id="AA07704.1"
/db_xref="GI:6382394"
/translation="MTRKTAIVARRKDAEICISGVRRKQENTNKLIVFLCLLE
GYLGVHRYVGKIGTGILYLFEGFLYGVLDLIRITNKKFCN"
710. .802
/gene="BB002"
710. .802
/gene="BB002"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA07738.1"
/db_xref="GI:6382428"
/translation="MRKGKIGICLVLEKSIIVGLMFKRIVR"
856. .1407
/gene="BB003"
856. .1407
/gene="BB003"
/note="similar to GB:L31615 GB:L31418 PID:520783
PID:551742 percent identity: 97.81; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein, putative"
/protein_id="AA07705.1"
/db_xref="GI:6382395"
/translation="MRLVGVFIIALALGCVLPDNOQAVOFEPFESSISMGSD
IVREGFSSKIVASEHRLIVEIKTILSKDPYRDPYRDPYNEFYKFLDGS
EÖSKDLILFGRYKNEONKKFSVYIWLSCISELSPDIKISGEESPEYDRMPAP
TAYÖQYLKVRKEIRENTINGKIL"
1404. .2268
/gene="BB004"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact: similar to
GP:520778 percent identity: 93.50; identified by sequence
similarity: putative; protein p23, authentic frameshift"
2744. .3496
/gene="BB005"
2744. .3496
/gene="BB005"
/note="similar to GP:1553115 percent identity: 53.30;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="antigen, p35, putative"
/protein_id="AA07706.1"
/db_xref="GI:6382396"
/translation="WKYICVCFLLINACNSDFSTNOEDIKYPSDEKSKSMEASS
KEEDPNKIRKNTLNDLINLEIANEHEKYEKKQEPSDÖYGIILFÖELDSVGKI
SEDPQSKFRKNTYSPLSAIDVKNLKLSEIINSQIOGLFNIIFNRGSIPEDSIN
HVSXKDILGGLTLDLKLNSPEKLSTKEPFSKMLNOLLDDYKNDKDHIRETNK
LKSHTLFPÖLDKKDEAYEPKNÖIFSISNL"
3623. .4108
/gene="BB006"
3623. .4108
/gene="BB006"
/note="similar to GP:2239281 percent identity: 79.75;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA07707.1"
/db_xref="GI:6382397"
/translation="MÖIYFÖEKLKGVGDIATVMKYESKRAITLAPLINTFSLIGIG
SFVGVYIGGALLGSOYLIGILIMAGYMGDTCFVSESTPATVITGCVLSIGTLTA
ASITGTGIIIPKFAKRYNADLKRRLGIALGELPNFDIGINGDSNRYLKNITETÖFTW
G"
complement(4249. .4986)
/gene="BB007"
complement(4249. .4986)
/gene="BB007"
/gene="BB007"

gene /note="similar to PID:1143400 percent identity: 62.50;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA07708.1"
/db_xref="GI:6382398"
/translation="MKKKYKKELVANSRFFSHMKELINIDMÖDKDILNNNLKEK
LKYNLDDIDNKIERMKILYIKKELFYDNSFPEÖTKEFVIKTAÖYALYKAYE
KYLEGASIDINKETGFRGLÖRQLEKNLNVKKNSEVNNKMSIRFPMKNEFYTE
CKEDTRACYILIERLEFTEKKEVSDLINYEKKYKKKKKKKYLIPYIGAYIILGMM
MSFVLPOLSSÖSKRPFMGFLFSFLLL"
complement(5069. .5830)
/gene="BB008"
complement(5069. .5830)
/gene="BB008"
/note="similar to GB:D26185 SP:P37522 PID:467381
PID:580906 GB:AL009126 percent identity: 54.17; identified
by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="plasmid partition protein, putative"
/protein_id="AA07709.1"
/db_xref="GI:6382399"
/translation="MDRKESNITIASPGVGKTTLTILFSYILKDSKKVLLIDLD
PÖNSLSYFSKYTFNDKNSYSLKKDYFEGCINKINDFISILRSHPLENSEI
LNYKDLLENILNRNTNVEDYILDTDPYFLKNSLNVYDIIIPÖVERSE
LSIILMÖTINDIKDIFNRKSPNISIIENÖFKKNRPFDELLYKESYFIKGIHY
NSVYLINELLESEVETEEETKNTLEINIINH"
complement(5803. .6339)
/gene="BB009"
complement(5803. .6339)
/gene="BB009"
/note="similar to GB:U03641 PID:458217 percent identity:
58.62; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA07710.1"
/db_xref="GI:6382400"
/translation="MNNFTEFRNLKKTKKRIFFYKIEKENKKTYSIIFKHLIK
EITRKGLRTEFOKENDGWNPFNLEPIKDEKELFGKMDNLKPEFLIGNNIKY
VVRKAYIEVFEKKSITKCYÖSLMTLLRREKKÖTYKKTLEINIMKEITYEFPNK
KIXEÖIINKWIERKNTL"
complement(6336. .6674)
/gene="BB010"
complement(6336. .6674)
/gene="BB010"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA07739.1"
/db_xref="GI:6382429"
/translation="MYNESEBAYALAVGPPALITPLFINKDPFKHLKYÖY
TKDKEFYLLKLLKYNKIHFMÖNHAHETDFFKIVGEFÖTITINKKIKNÖKTFLFD
LLANYKRRKI"
6800. .7411
/gene="BB012"
6800. .7411
/gene="BB012"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA07740.1"
/db_xref="GI:6382430"
/translation="MVRPMAIKGVQADATÖLLKDYÖDRRSKLEKFMKNPÖYDAGLL
SNSVEFRDKNIÖEFASGTSTKDFLENHFGYPPKRGIKRYIÖEKADÖIHIEPH
VEAGGEDLVGICIDIDEFSKATIVITNNFEGYLVAKSDTLKVKÖLIVFNRDGALE
KYKAPPNKATINAIASDAKÖISNDVYILIKAVFNGKAVSKN"

gene	7418..8392
gene	7418..8392
CDS	7418..8392
	/gene="BQ13"
	/note="similar to GP:1663561 percent identity: 100.00; identified by sequence similarity; putative"
	/codon_start=1
	/transl_table=11
	/product="conserved hypothetical protein"
	/protein_id="AA07711.1"
	/db_xref="GI:6382401"
	/classification="MGFENLEPENVYAKAVANIIGEVKDPIMYKWPSPQIEDVDLQMKORPVKMDPLINAPTLIANEVNTIIGSESEVRLNTYKLOKFKPHLKASRKFTYSDSLGIDINNLLPEFSQAYKLASSEIITKLINHPVLGTGVSIQKGKKNOKRLPPMA"
Query Match	13.8%; Score 104.8; DB 1; Length 52971;
Best Local Similarity	50.2%; Pred. No. 8.8e-09;
Matches 293; Conservative	0; Mismatches 282; Indels 9; Gaps 1;
QY	74 ATAAATTGAAAAAGTTAAATTTAAATACCTTTAAACCCTTTTAAATTTTCATTATA 133
DB	3473 ATATTTTCAATAGTAACTTTAAAGAAAAAGATTCACCTCATATATATTTAAATTT 3532
QY	134 TGTACCATAGTACAGTTTAAATAAAGGGCTTTTATGAAATCACTTTAATTTTGT 193
DB	3533 GTGTACATTCATTTATTAATAAAGCTATTTATAGAGAGTATATTAATATGAATAAT 3592
QY	194 TTGGCAACCTTTTGTGTTTTTTCTAGCTTTCCTCAAGCTAATGATTTCAAAAATGGTGC 253
DB	3593 TTCAATTAATATTAATTTTATAGTTTAAACAATGCAAACTTTGCACAGAGAACTTGA 3652
QY	254 TTTGGAGTGAAGTGTGCGGAAAAAAGCTTTGGTTTATGAAGTACGACAGCAAGATCTATT 313
DB	3653 AAAGGTGTGAGATATTCGAACCGTAAATGAATGAAGTAAAAAGCAACCATATTTA 3712
QY	314 GTACCATTTTAAAGAACCTTTTATAGGGTTGGAATAGGCTCCTTGCACAGAGAT 373
DB	3713 GCACCATTTACTTTGAAACATCTTTTATCTTAGAATAGGATCTTTGTTCAAGAGAT 3772
QY	374 ATTCTGGAGTTCCTTATCTCTTGAGTTGATCGGTTGATATAGGCTTATACCTGCG 433
DB	3773 TATATTTGGTGGGAGCACTACTGGCTCTCAAGTCCTTGGAGGAATACTTATATGCT 3832
QY	434 GGGGCTTATTTGGATATCAAAAGCGCTGATGATTTACTAAAAAAGCTTTTCATG 493
DB	3833 GG-----ATACATACCGGGGATTTGGATTCGTACGAAAGCAAGCAACATA 3883
QY	494 ACTTGCGGTAAAGGAGTATGTAGCAGGTGGTGTACTATGCGTGCACAGATTAACA 553
DB	3884 ATTAACCGAGAGATTTATTCAGGAATAGAGAGGCTCACAAATTCAGCAATCTACATCACT 3943
QY	554 GAAATTTATCTTCCATTTACATTTGCTATATGTTTAAATAGGAACCTAAAAATAGCTT 613
DB	3944 GGAATTTATCATTTCAATTTAAATTTGCTATATGATATCAACGACACCTTAAAAAAGACTC 4003
QY	614 AATGTAGCTTTAGAGAGATTTGAACCTAGTTTGTATGTTGCAT 657
DB	4004 GGCATTGCATTTGGCGGGGCTTGAACCAATTTTGACATTTGGAT 4047
RESULT 6	
LOCUS	AE000786
DEFINITION	29766 bp DNA linear BCT 17-DEC-1997
ACCESSION	AE000786
VERSION	AE000786.1
KEYWORDS	GI:2690008
SOURCE	Borrelia burgdorferi.
ORGANISM	Borrelia burgdorferi
REFERENCE	Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
AUTHORS	burgerferi group.
	1 (bases 1 to 29766)
	Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.,

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
burgdorferi	Nature 380 (6660), 580-586 (1997)	96065943	9403685	2 (bases 1 to 29766)	Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gattin, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J., Venter, J.C. et al.
burgdorferi	Nature 380 (6660), 580-586 (1997)	96065943	9403685	2 (bases 1 to 29766)	Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A., Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K., Gattin, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J.D., Weidman, J., Uterback, T., Matthey, L., McDonald, L., Attiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K., Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.
Direct Submission	Submitted (08-DEC-1997) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA			location/Qualifiers	
FEATURES				1..29766	/organism="Borrelia burgdorferi"
SOURCE				/db_xref="taxon:139"	/plasmid="lp28-2"
gene				116..1009	/gene="BBG01"
CDS				116..1009	/note="hypothetical protein; identified by Glimmer; putative"
				/codon_start=1	/transl_table=1
				/product="B. burgdorferi predicted coding region BBG01"	/protein_id="AAC66073.1"
				/db_xref="GI:26900307	/translation="MRSKSLFLYTLMLGLMSCLNLSKRLSSKEQKNNNVKEVSNSVQ EDGLNDLYSNQEKQSFTEKNGEKKWYEDLLNPPIIPSEPKKANIPNSIVHTORQ KEIKEDLIPSTNEKDEADEIKLEINELNSKFSELIREVLKDEYALINDPFDLQ VIEKHKKTKSIAMENYKNNRDKINKLNLNNLKINIELBOLIMIDIAENEISAAE FFPAOKRKLESITIRKLRKSNNRYALELSROLSPARSLSSLESAFRAAPPMVR KKRIKELIRHAKTYLESLNKK"
gene				1047..1928	/gene="BBG02"
CDS				1047..1928	/note="similar to PIR:A64689 percent identity: 35.78; identified by sequence similarity; putative"
				/codon_start=1	/transl_table=1
				/product="conserved hypothetical protein"	/protein_id="AAC66053.1"
				/db_xref="GI:26900107	/translation="MEINLSKLNNKNNKKLIFFISCSLVLVSTRPFDPNRTYSKRN GVIIIRPGYKIMKHLIEIQNNIALJTTTRLSKDRSNAFAVISISELSIIPGYVPYI IOESPEPRVKKENFKNFHFADLVAVNKFTAEELIGYVALIVSNYIDREYEHU QIDPPKIIFDINSIDFTLGKTGLDILNSHKVVPPTINIGSCFSPLDETCKNSGIS LEKYFEETRELNYNOTSRFINSKKEYVTNTISWTQLSKYLTRYRGRENPSKVEYEH LEKYKTIHYTHLTGIQKR"
gene				2104..2495	/gene="BBG03"
				/note="this region contains an authentic frame shift and is not the result of a sequencing artifact; similar to GP:2239281 percent identity: 63.64; identified by sequence similarity; putative; conserved hypothetical protein, authentic frame shift"	/complement(22750..2857) /gene="BBG04" complement(22750..2857) /gene="BBG04"
CDS					/note="hypothetical protein; identified by Glimmer;

```

gene
/putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG04"
/protein_id="AAC66074.1"
/db_xref="GI:2690031"
/translation="MKQSAISIFDEIAVHIKTLIDINLNLVINGVV"
/complement(2891..4056)
/gene="BBG05"
/Note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
GP:228714 percent identity: 99.12; identified by sequence
similarity; putative; transposase-like protein, authentic
frameshift"
4208..5368
/gene="BBG06"
4208..5368
/gene="BBG06"
/Note="similar to GB:U03641 PID:458216 percent identity:
66.76; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66054.1"
/db_xref="GI:2690011"
/translation="MKRVFTFLKLCITINIPRSMTINSSKPNCHNKLOQKLY
LSTLAVNSKYNKYTKQNLICENMLKRGQPTTLTQNLVLEKFKVTNY
YKHLNGEIEIYKLNKSKCYLKHNOHREKDYRFKADVDNYDENKGSYD
LVECLNNKNNIKERKIOIEKYOVAKYFKNCSLKEILPLINDINKDELKIL
KIIRKIELTKNNNTYLNYSFEKONKMLKILSNTOOLEKRGYMPDEISFOKI
YENKTPPHLEIENHKYDLSHKRLEKSEKKNLQDCENMRNINFEITLQOLK
KEYIDVLPKLIKLYNSKNKLEYNKVNMYELLEIIKKEKNOLKEVY"
5378..5950
/gene="BBG07"
5378..5950
/gene="BBG07"
/Note="similar to PID:114398 PID:1174342 PID:1174350
PID:2182755 percent identity: 57.84; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66055.1"
/db_xref="GI:2690012"
/translation="MQMMAKSILQKPIVRCNNKDLFIKIEKNDKTIYTKIMDI
YKFLGNKKNNKRYRISLGLFENOSKVEEFNLSLADFGIYGVYKRIKIKYKO
VGGIEKSYLSKTYMEVRFKKSIFCYFSLPILKKNVNTYNNALFMSMTLLER
OYGVDFYKXKYPQKGFPIKWIEMSLKNNHIL"
5911..6678
/gene="BBG08"
5911..6678
/gene="BBG08"
/Note="similar to PID:1174351 percent identity: 39.26;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="stage 0 sporulation protein J (spoJ)"
/protein_id="AAC66056.1"
/db_xref="GI:2690013"
/translation="MDRKLAKRPHIIAASIKGVGKSSIMESITLSTKNVULY
DLDPONAVTSYFITQDHPRMELINIVSYSLIKHKFKDVVISISKNLDPISYLEL
AKFESGNQREMLRNAVYNYEDYDIIDPPSSSELDNALVADKVIIPVPLE
RMAVENLPLINQIKELNNFMGKAKIHFASKSVIGRVTSTEISMLKEKYLNF
IGEVHSEALKKIDVAIGPRENENYKEYVRLIEKI"
6737..7288
/gene="BBG09"
6737..7288
/gene="BBG09"
/Note="similar to PID:1143400 percent identity: 43.68;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

```

```

/protein_id="AAC66057.1"
/db_xref="GI:2690014"
/translation="MKNRKVIINDIRVATYINNEERKKEELKNNELNRIEDD
IRKNTMKKILLEIRNKYLIIDCYKFEFDFIDRKARQAKYIKIATIDEGKLE
EIDITENGIDTKLENNKDKINSKANLITPLRLETOACDFYKMPNFANTILED
FYOKNEPOLIKLEENKOKYS"
/complement(7483..10779)
/gene="BBG10"
/complement(7483..10779)
/gene="BBG10"
/Note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG10"
/protein_id="AAC66075.1"
/db_xref="GI:2690032"
/translation="MKLDEITILPSMSISNNOKUIDAETLTKIADEKSSLDLKS
LEKSAGSNLEKALNKSQNAIKNFSLANSVNSVSKAGSKSIOGYLKSVDKIG
SVKMLANKTEAEFQMLAFAPIIMAVKAIQALISGTSIFDGMADLDFNEVSIF
SDMLGNDELKSLAESMRARFDETLFDRDAIANTKMLSYGATASEVERIRMGFA
AGSSSEGLEKLAEYSSVESNOVNLEDIALRDAGVDITDILAEERGLALYKRA
SDKIGSDALNKALSKATSSGCKFYGTAKERTIAAOOCTAKMSKILDLIGKALE
PMKIGPKYKQFLMVAIMVPLTKYTTITITLTKIGELVITLAKRYQKMKFKALT
TPYKILEGIRKIVYGLKLELATYVSSGFLDFRFGAEPILIKWQKLDIMTSKAYTLK
NLVFWKRAKREKEDTSGSEPERKGPDPNKTINMKMAEDYOKLDEIPNRDITN
KTGARQOALRNLEKTINENOKFKDEYSKIFPOLDENKKILVGEKSYNERNSNY
DEVNEYONLAKESRERETIKTLPHRDQVSALOKLNDEINENKMAFEVYKGFETL
NESRQVYNVLEKOVNEYETKLDRSVEVLELOKRTPLEMETMLPKREBASAEK
KMAKIDAMRKRYEYDEKRSQPKLINEINRTIQOYAKADDTKSLYSDMIDGLNFK
NAFKPDAVGFLENDGESIGEERHNILNKDRKARVWAGLDMOLESKTKITPA
GAVGPMGEAVALELNGLDVWGLILGQKSRKALAEKRDLDELEKRESELEK
LEDFFEDEIKMRKEKSELDEYTKSEIEFLQKOSKGOISGEERKQALDVQYETK
KDIETPOLTEEAKKLEVERKNNLSLEPERIAKQAEVDKVAEDMNMJGDEPRLK
TOKTIDELIRIARIKVSAGSIEETKILAHGARVSNKPYMPNSGVSSFRGQPELVR
ITPAIDENIRKLEAKITIAETITKLOQTSSIVNNFYTFNGDVLDAEKLVRMLSK
EHLMGFRMAE"
/complement(110776..11015)
/gene="BBG11"
/complement(110776..11015)
/gene="BBG11"
/Note="hypothetical protein; identified by Glimmer;
putative"

```

Query Match 12.2%; Score 92.6; DB 1; Length 29766;
Best Local Similarity 52.1%; Pred. No. 1.1e-06;
Matches 262; Conservative 0; Mismatches 229; Indels 12; Gaps 2;

```

165 TTTTATGATAAAGCTTTTAAATTTTGTGGACACTTTTGTGTTTCTAGCTTTG 224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2099 TTAATATGAAAAAATTTTCACTTAATATTAATTTTAAACAATGCAATCTTTG 2158
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 CTCAAGCTAATGATCTTAATAATGCGTGGAGTGGAGTGGAGAAAACCTTTGG 284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2159 CACAAGATAAAGCTTGAANA-----GGTGGAGTGGAGATATCGCGCTAATGA 2211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 TTTATGAACATAGACAGACAGATCCTATTTACACATTTTATGAACCTTTTAAAGGT 344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2212 AATATGAAGAAGGAAAAAAGCAACACATAGACACATCTTTGAAATATCTTTTGCTT 2271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
345 TTGGAATAGGCTCTTCTCTCAAGAGATATTTTGAAGGTTCTTATTTCTTGGATTGG 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2272 TAGGATAGGATCTTTTCAAGAGATATATATGAGTGAGTGGACATCTTGGCTCTC 2331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
405 ATGCGGTGATATAGGCTTATCTTACCTGGGGGCTTATTGGATATCAACGCTTGATG 464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2332 AAGTCTTGGA-----GGAATACTTAATATGCGTGGACACATGATCGTAGAGCTATGG 2386
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
465 GATATCTAATAAAGCTCTTTCAATGACTGGGATAGAGGATATTTAGACAGTGTG 524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2387 GAGCGTTACAGAAAGCACACACAGTAATATACCGAGGAGTATATACAGAAATAGAG 2446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
525 TGGTACTATGCGTGTGACAAGATTAAACGAATTAATCTTCATTTACATTTGCTATA 584
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

```

CDS
/gene="BBI08"
3576..3755
/feature="BBI08"
/feature="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBI08"
/protein_id="AAC6196.1"
/db_xref="GI:2690093"
/translation="MISFDPYNSASNIFIKITTPCKLRKLLPLIQKQFNNLS
FYKNIKSRSLNICK"
3745..3882
/gene="BBI09"
3745..3882
/feature="BBI09"
/feature="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBI09"
/protein_id="AAC6197.1"
/db_xref="GI:2690094"
/translation="MASFAKIDAEKSNKEVNSQKSKNNKNIEVKDTLGLVKAIRKS
L"
3911..4315
/gene="BBI10"
3911..4315
/feature="BBI10"
/feature="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBI10"
/protein_id="AAC6198.1"
/db_xref="GI:2690095"
/translation="MGRAMTPHQQLKYNLEGLKMINLDELKIAIDGNDNISKY
NDLKKVVDNENNSLIKRYLKNSNEDKLEAKKCIITLMHNETYESVCDLKDKN
KNEHNSILTLTLEIGIKNSAIAIHVCFNID"
complement(4623..4721)
/gene="BBI11"
complement(4623..4721)
/feature="BBI11"
/feature="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBI11"
/protein_id="AAC6199.1"
/db_xref="GI:2690096"
/translation="MSKLGYSYVNVYFPKIQIKLIYVNGSVLF"
5128..5346
/gene="BBI12"
5128..5346
/feature="BBI12"
/feature="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBI12"
/protein_id="AAC6200.1"
/db_xref="GI:2690097"
/translation="MIKPKNHISKIHIIKEVEDYTIKWDREYSLFKRLGKNKLTED
WLEYTOKEENOKIEFANKFKIKRKPRI"
5609..5707
/gene="BBI13"
5609..5707
/feature="BBI13"
CDS
/gene="BBI13"
5609..5707
/feature="BBI13"
Query Match 11.9%; Score 90.6; DB 1; Length 27323;
Best Local Similarity 50.4%; Pred. No. 2.5e-06;
Matches 325; Conservative 0; Mismatches 289; Indels 31; Gaps 3;

```

```

QY 24 TAATTTATAGTCATATCAGNAGCTCTATTGGAGCGAATTTCAACAAATTAATTTGA 83
Db 20256 TATATTTAAATCAAAATTAACCTTATTATACAACTCATCTAACCAAAATTTCA 20197
QY 84 AAAAGTTAAATTTAAATA-ACTTTAAAAACCTTTTAAATTCATTATATGTCACAT 142
Db 20196 GAAATTAAGATTAAATGATTTCTTGGGGTTTTCATTAATGCAATTAATTAATAT 20137
QY 143 AGTACAGATTAAATAAAGGGGTTTATGATAAATTAATTTATTTGTTGGCAAC 202
Db 20136 TAAATTTATTGGGGAATATATCAATATGAAAAAATTTTACATTTATTAATTTT 20077
QY 203 TTTTGTGTTTTTTCAGCTTTGCTCAGCTATGATTTAAATATGCTGGCTGGATG 262
Db 20076 AGCTTAACAATATCAATTTCTTGCACACAGATATGCTGAAAAAGTTGGAAGTAT 20017
QY 263 AGTGTGAGAAAACTTTGTTTATGAACTACAGCAAGATCTATTGTACCATTT 322
Db 20016 -----GAAACCATTTATGAATATTAAGCGAAAAAGCAACTATACACCATTC 19966
QY 323 TTATGAACCTTTTATGAGGTTTGAATAGGCTCTTGTCAAGAGATTTCTTGA 382
Db 19965 CTTTGAATTTATTTTATTTACTTTAGGAATAGGATCTTGTCCAAAGGATTTATTTG 19906
QY 383 GGTTCCTTATTTCTTGATTTGATGCGGTGCTATAGGCTTATACCTGCGGGCTTAT 442
Db 19905 GGTGCGCAGAGCTTGATCTGATCTAGTTATAGGAGGAATATCTTGATAGCTGAATAT 19846
QY 443 TTGATATCAAAAGCGCTTGATGATTTATCTAATAAAGCTGCTTTCATGACTGGGGT 502
Db 19845 CTTGGCATPACGATGATGAACCAAGACCAACACTG-----GG 19807
QY 503 AAGGAGTTATGTTAGCAGGTGCTGTACTATGCTGTGACAGATTAACAGAAATTA 562
Db 19806 CATATATATACACAGATAGGATAGGATAGCAGATATAGCATCCCATATGCTCATTTAT 19747
QY 563 CTTCCATTTACATTTGCTATATATGTTATATAGGAAGCTAAAAATATAGCTTAATG 622
Db 19746 ATTCCATTTACATTTGCAAAATTAACACATGCAATCTTAAAAAAGACTGGCATGAT 19687
QY 623 TTAGGAGATTTGACCTAGTTTATGATGTCGAATGGCCCAATCC 667
Db 19686 ATTCGGGTTTGAACCCCAATTTGATATGGAATAGCGGATTC 19642

RESULT 8
AC119697/c 173540 bp DNA linear HTG 18-JUL-2002
LOCUS AC119697
DEFINITION Rattus norvegicus clone CH230-135B13, *** SEQUENCING IN PROGRESS
***, 61 unordered pieces.
ACCESSION AC119697
VERSION AC119697.4 GI:21747171
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 173540)
AUTHORS Alzbrock,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Muzny,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieve,M., Brown,E., Brown,M., Bryant,N.P.,
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
David,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dihn,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

```


Correll, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Hawlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Huliy, K., Hume, J., Jackson, L. E.,
Jacobson, B., Jia, Y., Johnson, R., Jollivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maneshwar, M., Mapua, P., Martin, R., Martindale, A., Matlhez, E.,
Massey, E., Mathney, E., McLeod, M. P., Meador, M., Mei, G., Metker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenko, S., Oguh, M., Okwuonu, S.,
Oregunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pui, L. L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Swatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umanil, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gids, R.

Direct Submission
Unpublished
2 (bases 1 to 173540)
Worley, K. C.

Direct Submission
Submitted (01-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 173540)
Worley, K. C.

Direct Submission
Submitted (18-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20467833.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVPY
Center clone name: CH230-135B13
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 113535 bases at least Q40
Consensus quality: 119717 bases at least Q30
Consensus quality: 124921 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 61 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1043: contig of 1043 bp in length
*	1044
*	1144: gap of unknown length
*	1144
*	2163: contig of 1019 bp in length
*	2163
*	2262: gap of unknown length
*	2263
*	4101: contig of 1839 bp in length
*	4102
*	gap of unknown length

```

* 76922 77021: gap of unknown length
* 77022 79872: contig of 2851 bp in length
* 79872 79972: gap of unknown length
* 79972 83509: contig of 3537 bp in length
* 83510 83609: gap of unknown length
* 83610 86201: contig of 2592 bp in length
* 86202 86302: gap of unknown length
* 86302 88570: contig of 2269 bp in length
* 88571 88670: gap of unknown length
* 88671 92159: contig of 3489 bp in length
* 92160 92259: gap of unknown length
* 92260 95582: contig of 3323 bp in length
* 95583 95682: gap of unknown length
* 95683 98902: contig of 3220 bp in length
* 98903 103202: contig of 4200 bp in length
* 103203 103302: gap of unknown length
* 103303 107401: contig of 4099 bp in length
* 107402 107501: gap of unknown length
* 107502 112766: contig of 5265 bp in length
* 112767 112867: gap of unknown length
* 112867 11657: contig of 3791 bp in length
* 11658 116757: gap of unknown length
* 116758 120996: contig of 4239 bp in length
* 120997 121096: gap of unknown length
* 121097 124159: contig of 3063 bp in length
* 124160 124259: gap of unknown length
* 124260 128734: contig of 4475 bp in length

```

Query Match 8.98; Score 67.6; DB 2; Length 173540;
 Best Local Similarity 45.58; Pred. No. 0.01;
 Matches 240; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

```

QY 151 TTTTAATAAGGGGTTTATGATTAACCTTTATTTTGGACACCTTTGTGT 210
Db 12741 TGTATTTCTCTGTTGTTTGTATTCCTGTTGTTATTTGTTGCTGTTGCTGT 12682
QY 211 TTTTTCAGCTTGTCTCAACGTAATGATTTCTAAAAATGCGTTGGATGATGCTGG 270
Db 12681 TGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12622
QY 271 AGAAAACTTTGTTTGAACAGCAACAGATCTTATGACATTTTATTGAA 330
Db 12621 TGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12562
QY 331 CCTTTTAAAGGTTTGAATAGGCTCTTGTCTCAAGAGATTTCTTGGAGTTCTCT 390
Db 12561 TATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12502
QY 391 TATTTCTGATTTGATGCGGTTGATAGGCTTATACCTTGCAGGCTTATTTGGATAT 450
Db 12501 TGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12442
QY 451 CAAAGCGCTTATGATGATTAACAAAAAGCTCTTTTCAATGACCTTGGGTAAGGAGT 510
Db 12441 TGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12382
QY 511 TATGTTGACAGTGTGTTACTATGCGCTGACAAAGATTACAGAAATTATTTCTCAT 570
Db 12381 TGTGTTGTTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12322
QY 571 TACATTTGCTAATGATTAATATAGGAAGCTAAAAATAGCCTTAATGACCTTTAGAGAG 630
Db 12321 TATGTTGCTAATGATTTTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12262
QY 631 ATTGAACCTAGTTTGTGATGTGCAATGGCCCAATCCAGTCTCTCTG 677
Db 12261 TGTGTTGCTGCTGCTAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 12215

```

RESULT 9
 AE000784/c AE000784 28601 bp DNA linear BCT 17-DEC-1997
 LOCUS

```

DEFINITION Borrelia burgdorferi plasmid lp28-3, complete plasmid sequence.
ACCESSION AE000784
VERSION AE000784.1 GI:2690041
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 28601)
Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.,
Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K.,
Gavin,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D.,
Richardson,D., Peterson,J., Kierlavage,A.R., Quackenbush,J.,
Salzberg,S., Hanson,M., Van-Vugt,R., Palmer,N., Adams,M.D.,
Gocayne,J.D., Weidman,J., Uterback,T., Matthey,L., McDonald,L.,
Artlich,P., Bowman,C., Garland,S., Fujii,C., Cotton,M.D., Horst,K.,
Roberts,K., Hatch,B., Smith,H.O. and Venter,J.C.
Direct Submission
Submitted (08-DEC-1997) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
location/Qualifiers
1. 28601
/organism="Borrelia burgdorferi"
/db_xref="taxon:139"
/plasmid="lp28-3"
273..467
/gene="BBH01"
273..467
/gene="BBH01"
/note="similar to GP:1655789 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="A065980.1"
/db_xref="GI:2690046"
/transl_table="MNRLLIFLLVFSNLSNSDNNPLMSNKEKISEYQINSSN
KSYIFKRNSSVKKRYTNHYY"
391..858
/gene="BBH02"
391..858
/gene="BBH02"
/note="similar to GP:1655790 percent identity: 97.44;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="A065992.1"
/db_xref="GI:2690048"
/transl_table="MSRQTNIGFSNEIOALKDRSTIITNQNNDINSTINYPYIOTI
LKEIKVDGNGINGTKESGETKLEILNGNISREKDAIOYGSGPFAKDVRENQOK
ENKMSHIHVDDEKREYHILMPSINNADSSSYITNYIINDNLRLISNL"
926..1075
/gene="BBH03"
926..1075
/gene="BBH03"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH03"

```

```

gene
/Protein_id="AAC65997.1"
/db_xref="GI:2690053"
/translation="MKKNKKCEIEKTEQLEIINNQSEIKOLHOLEIETTCCLLY
VTIH"
/gene="BBH04"
1045..1368
/gene="BBH04"
1045..1368
/note="Similar to PID:1142590 GB:AE000793 GB:AE000784
percent identity: 38.46; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAC65986.1"
/db_xref="GI:2690042"
/translation="MPALCDWTLNLNLEINYSQKRLKLYNEILKKDKNSCDLEPTN
KYDILEKTKTIVKLSFRNOSKMYIKYINPLVFRSTIDYQYQITDKLRLLEIN
YPTI"
1498..1680
/gene="BBH05"
1498..1680
/note="Similar to PID:1142590 GB:AE000793 GB:AE000784
percent identity: 93.88; identified by sequence
similarity: putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAC65991.1"
/db_xref="GI:2690047"
/translation="MNSKNRLSHLILYTLNNDLNSKNIGYSGFIRAFTEFI
DKVSKLNYFLYFPIK"
complement(2260..2970)
/gene="BBH06"
complement(2260..2970)
/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH06"
/protein_id="AAC65998.1"
/db_xref="GI:2690054"
/translation="MKRSFLSYMLISLISCSVSRINQINNELKTFYKAYYSI
KLDAIYNECTATYNDIMYISBGTSDSKVQASIFEKDKIVNKEFELIIEYK
PMFLSKLIDFAILEDVNDVSNARHVADSYKRLKRSVLAYIESDVISKFDVS
KFEVASKRFVYKAKAFVEENDLIALECIVKTIQDMVNDREINSRSRYNNFYKKEADFL
GAAYELGAYKAIKOTLL"
complement(3083..3514)
/gene="BBH07"
complement(3083..3514)
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact.; similar to
GB:003641 PID:458217 percent identity: 63.33; identified
by sequence similarity: putative; conserved hypothetical
protein, authentic frameshift"
complement(3590..3730)
/gene="BBH08"
complement(3590..3730)
/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH08"
/protein_id="AAC65999.1"
/db_xref="GI:2690055"
/translation="MGVFIKAKTLEIKSFLSKRRSSGCHISILLYAKEDDLAI
NM"
complement(3892..7728)
/gene="BBH09"
complement(3892..7728)
/gene="BBH09"

```

```

/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH09"
/protein_id="AAC66000.1"
/db_xref="GI:2690056"
/translation="MVVMNNNSRFIITNDPNNSLYKELSGFKIKENIYKSNFPI
FLKNQIAIDNSTEANTESLKLSTFEELASVEQKGGQIEGVSRDILLFENDKD
KAFNRKLEAKKNNEISPTEDDILIAVKKRPSFEDKDKVEADQLYRLNQQYK
HAGLSNGKVRKYDYSKVLGEKRYIEFDSKIKEEYKEQEMFLLFSYLKERY
LKTSNIISVEKEQISKEKEIIOKTLRELYRPPDSIYKIAKNYKDEFPDENKEIT
PHILASILEESIPIRLREFIAYIEDNDIFKLTLEIYSSVSFPFEPDENKEIT
LGKKIITIFMLDKGSAIKRPIFNGLPQADKKYIANNESLISISIEELKYLIF
FEENIKDKKTFVKSRLDKSFGELYETLELDLADTAYHRIVEDSILYLRHEEL
ENNRVKNATYALIKNIYLYTSRLDKMGAAVYTPDDLDLFWVISIEQDLKSPLDI
KELIDNSCGSHFLISCLDYLETKEVWEELDKEDVAKELDEKYGILLKSEYDIQDSI
SKEIYAKMLLRCIYGVADINPISVEITMLSMINTFEGPLSPFIEHHIKGNALLG
YTKDEFPDIKKKPFESGSLFKRKRIKEITTLLEDGYOKIKGINDTKEDIRSKKIYK
EYESKYDNRIRIIEPLIKLISDPSKLNIEFDQIAVVISLIEYIKNKSISDNEK
MEKIRLSHKKFRHYGIEFPDIOEGFIVYGNPWEKTKNEAFESKHIPSTRKLS
IKQONKIKOETLGRDHPNLNIEYNEKSMGTINNLKSDKRNFSAGDPMLFRYFA
ENLKLKENGMLTYLVPALMSSESSARILRYIFTNYKLANIYOFONOKRKYDATTLE
KFAIFQISNTKTPSNEFKAKEMIOSNDILKEITSNLENKDDPYKGIKIDIOIKL
SPIQESIIEFDSKELILINKESKFSYLSSEYINFGVGLMUTYKTLCKYNNENFI
FLYSGANHOFNSRFEEDKAKESKILMDKDDPOKYSTONQOYIEKILYRLIARN
TDIRMIISTLSPRNCYCVSTIYINTEKPIYSIYKKLPIISIFNSVPDYLIARRALST
DIVASCIQCPMPQPEEELIAPLYLVLNVTSLILAKNOLPNSLTYLEHFRFSK
EKVNIKLKLDTKDEFKKEENENNEIVASLYSLKDEFEVLNDFEVLNKKNGEYIL
FLIKGYENLYKTKNLN"
complement(8000..8203)
/gene="BBH10"
complement(8000..8203)
/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH10"
/protein_id="AAC66001.1"
/db_xref="GI:2690057"
/translation="MNKNVILICVFDLISFKNFLHLCJLHSDQVDFKEIKRLY
NCFNEKDKNIPLFMLHKLGNV"
complement(8701..8796)
/gene="BBH11"
complement(8701..8796)
/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH11"
/protein_id="AAC66002.1"

```

Query Match 8.7%; Score 66; DB 1; Length 28601;
Best Local Similarity 52.1%; Pred. No. 0.033;
Matches 221; Conservative 0; Mismatches 170; Indels 33; Gaps 2;

```

gene
complement(8701..8796)
/gene="BBH11"
complement(8701..8796)
/note="hypothetical protein; identified by Glimmer:
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BBH11"
/protein_id="AAC66002.1"

```

304 AGATCTATTTGACATTTTATGAACCTTTTATGAGGTTTGAAGATAGCCTCTTGC 363
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27997 AGATCTATGATTCAGATATGTTAAATTTATTTTACCTTTTGGAGTGGGCTTTTGT 27998
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27937 CCAAGAGATTTATTTGAGGCTGACACGCTTGATTTATTTGATGATTAACCT 27878
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27877 TTTAACAACTGGATTTCAATGAATAAATCTGAAAAAGAACCTGATGATATCTTC 27818
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 484 TTTTCATGACCTTGAGGATTAATGACAGGTGCTTACTATGAGCTGATC 543
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 27817 GAT-----GATCTACTACTTTCAGGAATGCTACAGTTTGGATCATC 27776

QY 544 AAGATTACAGAAATTTATTCATTTACATTTGCTTAATAGTATTAATAGGAAGCTTAA 603
 Db 27775 CTACTTATATCAATTTATTTGGCCAGTACTATTAAAGATGATCTACAAAAATCTAT 27716
 QY 604 AAATAGCCTTAATGAGCTTTAGAGAGATTTGAACCTTATTTGATGTCGAATGGCCCA 663
 Db 27715 GAATCGAATCATTTGATGAGACTTGGCGGATTTGAACCAATCTTGACATTTGAATGAAC-- 27658
 QY 664 ATTCAGTCTCTGGCTTGAAGTCTTTCAAAAAAGCATTAATTTTATTTTACA 723
 Db 27657 -----GGATTCACAATGCTCCTTTAAAAAAGTTATTAATTAATGATATAA 27611
 QY 724 AAAA 727
 Db 27610 GATA 27607

RESULT 10
 Locus 1592044/c
 DEFINITION Human DNA sequence from clone RP13-407F1 on chromosome 13, complete sequence.
 ACCESSION AL592044 GI:16151294
 VERSION AL592044
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 236130)
 AUTHORS Pearce,A.
 TITLE Direct Submission
 JOURNAL Submitted (12-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquereyesanger.ac.uk
 COMMENT requests: clonequest@sanger.ac.uk
 On Oct 13, 2001 this sequence version replaced gi:14530013.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WormPeP; Information on the WormPeP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr13>
 RP13-407F1 is from the library RPCI-13.2 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/Dacpac/home.htm>
 VECTOR: pBACe3.6
 This sequence is the entire insert of clone RP13-407F1.
 FEATURES
 source
 1..236130
 location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /clone="RP13-407F1"
 /clone_id="RPCI-13.2"
 BASE COUNT 79573 a 39654 c 39207 g 77696 t
 ORIGIN

Query Match 8.2%; Score 62; DB 9; Length 236130;
 Best Local Similarity 62.6%; Pred. No. 0.09;
 Matches 114; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

QY 70 AGCAATAATTTGAAAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 129
 Db 75626 AGAATCATATGAAAAAATATATATCTTTAAATACATTTGAAAAACCTTAAAAAATTTAT 75567
 QY 130 AATATGTCACATAGTACCACTTTTAAATGAAGGGTTTATTAATGAATTAATTAATTT 189
 Db 75566 CAAATGATTAAGTACGTTTCACTTTTATTAAGCAGATTAATTAACATTAATTAAT 75507
 QY 190 TGTTTTGGCAACCTT---TGTGTTTTTCTAGCTTTGCTCAACGTAATGATCTTAAAAA 246
 Db 75506 TGTATGAGATTTTATATGTGTGATTTTTTTCACATATGAACCTTAGTATTAATTA 75447
 QY 247 TG 248
 Db 75446 TG 75445

RESULT 11
 Locus AC117080
 DEFINITION Dictyostelium discoideum chromosome 2 map 780304-886736 strain AX4, complete sequence.
 ACCESSION AC117080 GI:20066270
 VERSION AC117080
 KEYWORDS HTG.
 SOURCE Dictyostelium discoideum.
 ORGANISM Dictyostelium discoideum.
 Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
 REFERENCE 1 (bases 1 to 106434)
 AUTHORS Gloeckner,G., Eichinger,L., Szafanski,K., Pachbat,J., Dear,P., Lehmann,R., Baumgart,C., Parra,G., April,J.F., Gulgo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and Noegel,A.A.
 TITLE Sequence and Analysis of Chromosome 2 of Dictyostelium
 JOURNAL Unpublished
 REMARK The Dictyostelium Genome Sequencing Consortium
 REFERENCE 2 (bases 1 to 106434)
 AUTHORS Baumgart,C.
 TITLE Direct Submission
 JOURNAL Submitted (06-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
 COMMENT is available from IMB Jena, Department of Genome Analysis (<http://genome.imb-jena.de/dictyostelium/>) and the University Cologne, Institute for Biochemistry I (<http://www.uni-koeln.de/dictyostelium/project.shtml>)
 Funding Agency : Deutsche Forschungsgemeinschaft (DFG).
 FEATURES
 source
 1..106434
 location/Qualifiers
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /chromosome="2"
 /map="780304-886736"
 175..1003
 /note="ORF_id:dd_00962"
 /pseudo
 /codon_start=1
 complement(join(1326..1349,1442..1643,1999..2203,2239..2246))
 /note="ORF_id:dd_00960"
 /pseudo
 CDS

```
/codon_start=1
/complement(join(2301..2392,2533..2590))
/note="orf_id:dd_00996"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45297.1"
/db_xref="GI:21281410"
/translation="MELEKVERKMLELEVGDYVDYFDIPSELSISFGNESSWLNTSDKF
KTFPLN"
complement(join(3047..3175,3362..3927,4003..5929,
6078..6206))
/note="orf_id:dd_00995"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45298.1"
/db_xref="GI:21281411"
/translation="MSDNEQESSQVVLKEDKTVTFVFQSLVSSNEDTDTLRDRK
SLKWSISDPNPKRKIKIDNGSLTTAASSQOQOELGLAVLIRQYEENIVKEL
LDEKKIETNSMKPNKQOMELIKKSGPNTQMFEDVLNGNCEGVMALKVTREK
SIVGIFSGDATFTFTIGVSCFMDNDNLSNLSFTIMQSVKELCCDQKNTDYQVKE
KLSDAGIDPTELPRKSDSSKABODIPRLGSKVKNLPIDIQEHAIQSASCLIKLID
LSNPTNREKFKLEYDIDRIRKLDSSSKGLHITDLKDSVSAAGGAGAGASSSN
KDQSLYNLQNCNTPMGSRLLQWVKQPLNABELEARNFEVAYNDLELRQSLRN
DLKRTIGDLRLSKRHGOKATLEDCVNLGYVTRLPVYLSLNHNSIHOELIKVNF
ESLESISDFPAKPCAMVEKTDLDLANDKHEVIRSEFDLRLQKDOISNKIER
FRVADIADLNDEAKVLYSEKDMELLRIKRDENVAIRDKKRYVHATVADQVFEAT
REIDTLNFAVKKMSAEYLDKODGLAKRTLQIATSFVPLIEDLSLITLDFVITSHV
SSIAVPIRPEIIPGSDENGAGVLIIGGRHPCVEIDNVNFLANDIDLTRGQSOFQ
IITGPMGKSTFIRQVGLIVLMAOIGCFPAOKATIAVYQCIISRGAGDSOLRGVS
TFMAEMLETSYLIVATKNSLIIDELRGSTSGFGLWAGIAYICNOIGGFLFA
THFHELTISDLPLPVKRLHYASASTQNTFTLLYKVEGPGDQSGIHVALIANTPSO
VIENAKQAKLESESESWTLKONHNKFLFEKKEINFNENDVKEKSLVYNSLTKYSD
IN"
complement(join(6582..6743,7143..7178))
/note="orf_id:dd_03346"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45299.1"
/db_xref="GI:21281412"
/translation="WLSGLKIDKEKFFLIIIIINNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNND"
8201..9358
/note="orf_id:dd_00993"
/codon_start=1
/product="5'-adenosylmethionine decarboxylase proenzyme"
/protein_id="AAM45300.1"
/db_xref="GI:21281413"
/translation="MENNNDYKRNSTEIFNEADAGVYQSNNDYSDGEGPEKKLID
RGPISKGSVKSIAGPSKVGIRITIDREKQWTVLDSARCTIISQSNHMSYVSES
SLFVPRAMIKTCTTLLHLIAKMOVGKEGCEVEYVFSKRLNLPQSKQVPHC
SFSDBVNLINKIFDQATVWGDVNDKHNKLIADFRKNPTLQRTQETVEVNHIDET
VMKQFKEGVSAMDVTNCGIADLPGSMIDDFQDPCGSMGLNEPFWHTITTP
ESHSYVSFTDNVALADYNOLAKLVANFKGRTAALYAEAGACGDPYAFDVNP
SYAIONKTVHGEFGDGYVANSYQOLDKKVNSNDLQGISLSTNIDLIQY"
complement(9773..10825)
/note="orf_id:dd_00992"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45301.1"
/db_xref="GI:21281414"
/translation="MKCFILNNIKRKNTPNNIFKASYSHPFNNSNNIIRNFSSNKI
NENSNNKEDNMLIPSRSLRSLKESLSLAYLNEEYTEEVEDEEDDDNNN
NNNNNEIKEEKEEKEYNQNLNEKGIPLKGGKRFQDMYRIKENHELSNKSITLND
VGEADSEIEIRKVEERLKKRYOEKMEILDSMDSQDFYKDKNEKVDLKNISENKA
DTHETPEITDKOLENNOLKPNGLSISMSDSAPFNKYQOSDSSRISKIRIO
KDKLNRHINGENLVDFSLYREDPILNNAQFLAKKFLMSKSVENLKYTIPIIVK
FNGIKSGHNVAYEER"
join(12150..12227,12323..15037)
/note="orf_id:dd_03068"
/codon_start=1
/product="unknown protein similar to ESTs emb|246294,
gb|W43468, and gb|W43475"
/protein_id="AAM45302.1"
CDS

CDS
/complement(join(2301..2392,2533..2590))
/note="orf_id:dd_00996"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45297.1"
/db_xref="GI:21281410"
/translation="MELEKVERKMLELEVGDYVDYFDIPSELSISFGNESSWLNTSDKF
KTFPLN"
complement(join(3047..3175,3362..3927,4003..5929,
6078..6206))
/note="orf_id:dd_00995"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45298.1"
/db_xref="GI:21281411"
/translation="MSDNEQESSQVVLKEDKTVTFVFQSLVSSNEDTDTLRDRK
SLKWSISDPNPKRKIKIDNGSLTTAASSQOQOELGLAVLIRQYEENIVKEL
LDEKKIETNSMKPNKQOMELIKKSGPNTQMFEDVLNGNCEGVMALKVTREK
SIVGIFSGDATFTFTIGVSCFMDNDNLSNLSFTIMQSVKELCCDQKNTDYQVKE
KLSDAGIDPTELPRKSDSSKABODIPRLGSKVKNLPIDIQEHAIQSASCLIKLID
LSNPTNREKFKLEYDIDRIRKLDSSSKGLHITDLKDSVSAAGGAGAGASSSN
KDQSLYNLQNCNTPMGSRLLQWVKQPLNABELEARNFEVAYNDLELRQSLRN
DLKRTIGDLRLSKRHGOKATLEDCVNLGYVTRLPVYLSLNHNSIHOELIKVNF
ESLESISDFPAKPCAMVEKTDLDLANDKHEVIRSEFDLRLQKDOISNKIER
FRVADIADLNDEAKVLYSEKDMELLRIKRDENVAIRDKKRYVHATVADQVFEAT
REIDTLNFAVKKMSAEYLDKODGLAKRTLQIATSFVPLIEDLSLITLDFVITSHV
SSIAVPIRPEIIPGSDENGAGVLIIGGRHPCVEIDNVNFLANDIDLTRGQSOFQ
IITGPMGKSTFIRQVGLIVLMAOIGCFPAOKATIAVYQCIISRGAGDSOLRGVS
TFMAEMLETSYLIVATKNSLIIDELRGSTSGFGLWAGIAYICNOIGGFLFA
THFHELTISDLPLPVKRLHYASASTQNTFTLLYKVEGPGDQSGIHVALIANTPSO
VIENAKQAKLESESESWTLKONHNKFLFEKKEINFNENDVKEKSLVYNSLTKYSD
IN"
complement(join(6582..6743,7143..7178))
/note="orf_id:dd_03346"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45299.1"
/db_xref="GI:21281412"
/translation="WLSGLKIDKEKFFLIIIIINNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNND"
8201..9358
/note="orf_id:dd_00993"
/codon_start=1
/product="5'-adenosylmethionine decarboxylase proenzyme"
/protein_id="AAM45300.1"
/db_xref="GI:21281413"
/translation="MENNNDYKRNSTEIFNEADAGVYQSNNDYSDGEGPEKKLID
RGPISKGSVKSIAGPSKVGIRITIDREKQWTVLDSARCTIISQSNHMSYVSES
SLFVPRAMIKTCTTLLHLIAKMOVGKEGCEVEYVFSKRLNLPQSKQVPHC
SFSDBVNLINKIFDQATVWGDVNDKHNKLIADFRKNPTLQRTQETVEVNHIDET
VMKQFKEGVSAMDVTNCGIADLPGSMIDDFQDPCGSMGLNEPFWHTITTP
ESHSYVSFTDNVALADYNOLAKLVANFKGRTAALYAEAGACGDPYAFDVNP
SYAIONKTVHGEFGDGYVANSYQOLDKKVNSNDLQGISLSTNIDLIQY"
complement(9773..10825)
/note="orf_id:dd_00992"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45301.1"
/db_xref="GI:21281414"
/translation="MKCFILNNIKRKNTPNNIFKASYSHPFNNSNNIIRNFSSNKI
NENSNNKEDNMLIPSRSLRSLKESLSLAYLNEEYTEEVEDEEDDDNNN
NNNNNEIKEEKEEKEYNQNLNEKGIPLKGGKRFQDMYRIKENHELSNKSITLND
VGEADSEIEIRKVEERLKKRYOEKMEILDSMDSQDFYKDKNEKVDLKNISENKA
DTHETPEITDKOLENNOLKPNGLSISMSDSAPFNKYQOSDSSRISKIRIO
KDKLNRHINGENLVDFSLYREDPILNNAQFLAKKFLMSKSVENLKYTIPIIVK
FNGIKSGHNVAYEER"
join(12150..12227,12323..15037)
/note="orf_id:dd_03068"
/codon_start=1
/product="unknown protein similar to ESTs emb|246294,
gb|W43468, and gb|W43475"
/protein_id="AAM45302.1"
CDS

CDS
/db_xref="GI:21281415"
/translation="MDKYDTILMLRLRELEENSKCADCTDSFPRVYNTTYGVFVCSVCG
AHLRELGNRVYSISDKRTQODIERLEKVGKMADEITLSSWQOQVLPSPSEKRY
KDFIKMKYTEKWKIEGITSQDFISPPQSTPGSTPTSTRQOQOQOQPSAVSQKQQT
TSPLRNSEKKEIILTSKDSLSEKESESLGSLNLLNTNQOQOQOQOQOQOQOQOQPSQ
TSQOISPEIIOAEINQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
SEKPEQNSYLNKSTYMAADSLGNSNSFYKIQVQHPPKSPSISIDPTQOYKQOQ
QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
NOMGHANOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOMNOM
NOLALVPEQNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
INMNSNSNINPENGPEGISDQVYLKHQOHQOQNNNNNNNNNNNNNNNNNNNNNNNN
PATTTTTTIHTPIVGNLPSSESSSSSGNPFQOQOQOQOQOQOQOQOQOQOQOQOQ
SSFEESGNSIDRDEDEKKAEMRIQEKNIQIONDEIARKILOEEELFKTRPSRTE
PSPRRTSOLQOQOPOQOQOPOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
SSGNILKRLSGRSRKSYDDDPPEINPRLNPGSGYDQOQOQOQOQOQOQOQOQOQOQOQ
OPKISQDITYODNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
YCGKQVGLSESSYHAKNTECILRQERCSMCKLVKHMQGDHQLQCTVTVTCVGCGR
VMGKEMQHMAVYHNO"
join(15592..15843,15949..16261,16362..16444)
/note="orf_id:dd_00988"
/codon_start=1
/product="SIMILAR TO S. CEREVISIAE NUCLEAR PROTEIN SNF7"
/protein_id="AAM45303.1"
/db_xref="GI:21281416"
/translation="MKLRGPKPKODTOETIGKRLTLDMLEKROTFLQDLKEQEK
AKVLYSQRKREALLCKLRNNRQNEVTKLGQSYDTLQOQIFALENAKMNNEIIMSRR
DGARSRLHGLHITDKYDDVYIEEIOEOMETHEBELSNAISQPLQVQVDEBELLELA
EYDEBELDAQLLNTKAPSRLEPQETQVMEPIYSKTVPLSKSEEEIRLESBLAM"
complement(16713..18956)
/note="orf_id:dd_03336"
/codon_start=1
/product="PUTATIVE MITOCHONDRIAL RNA SPLICING PROTEIN"
/protein_id="AAM45304.1"
/db_xref="GI:21281417"
/translation="MNKLINYLKCKSNKLIFFPNNOQRYFTKYNLSTPISPLNK
EDNNDYDENNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
VSTSTIKTFDESEFSKPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
RHYKSLSNELKQARLDLRTDPSPPMPTLVNAYILISIGCVARIVQYRVNPLV
DGNQTLIDETAGIHESLTSQGTETLPPPEFKVESILDLCKLEFBERMOSLI
EKELOMLNENPEHNLIELLYHKKKLNQFVKIKIIDAITDLLEADMDALMYLSFR
HATGARKNQHDSEILEETVROLLELSFQISQPLKSTGEFVFOJDTARNMK
MRMNLMLSVTISGLGSVIAAGTFGMEISQPLAFPIAGCSIAICIGLFTFGIK
YYCHVKNTLPTYKKTPTPALGALGFRNSDIPDNLILKEQSOHJILKEHOSLFT
NNTNTNTNTNSNTFNSTNDSSKANNIOYHKHIDPENOQITDOONKRIQVKNIQOF
QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ
LTTNENSTNDSTIILEDTLNSSLNNQNNQNNNNNNNNNNNNNNNNNNNNNNNNNN
KHOSQYHNPOLAPNLISFWKVIQESVYNSDDTTRNNININNNNNNNNNNSPFLSS
TSNRNL"
19965..20948
/note="orf_id:dd_03337"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAM45305.1"
/db_xref="GI:21281418"
/translation="MNSFKNTIYVLIITFFLNKIKSEICKNOLLAYFENNNNNN
FLSLELSTIILKNNQIITYGNEHLKEBEKNLVNDGLIGIONDKIPVLSNDPESDS
PNSLNRTRVYVYGEFYIHKDGLIIGSCGNKLEPIILCNDITFSLGLDDLIQDQIKT
GSNNNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
GSSNNITKYSLVKIVILNNSNIIINNINNINNINNINNINNINNINNINNINNINNS
GVNKYISPLNNNNTTIFSLPSYSLPISPGKYEGEGLIKGSTTPIRKIP"
join(21340..21871,21963..22864,22954..23622)
/note="orf_id:dd_00979"
/codon_start=1
/product="PEPSTATIN-INSENSITIVE CARBOXYL PROTEINASE 2"
/protein_id="AAM45306.1"
/db_xref="GI:21281419"
/translation="MKNINLPIILILINCYRGINQVNGKPHQELGKQOSLTFEIPD
NPTTFRIYVQKNEQLKIKRVDSYDPSHNSNGVLSYHQTGKILISPSKESLTIK
WLOPNPIEYNDQDMETITMSKONIKEMLNINIFTHQNGKILIRSPEDLPIPS
NLKIHIDITIGISHPPYVKKKINSLSSSSSSSSGEGLQIISIKCTGEBFLYVQPS
SPNCGKYEPMQVTVKALVQNNNTITNSVNPSCVNDGAVGVKCVANSFVQYQSE
STIDPESNOVLMDHPIVSLPIVLPOTIKOYGINNHQVSPHNPATOCVVEEEOQYS
PTDLEQFNSMGLPPIVNSVNYGVNDEINPGEASLDIOLKMGVAPMNOTTFWMSIYAN
```


RESULT 14
AC121387/c
LOCUS
DEFINITION AC121387 193379 bp DNA linear HTG 22-JUL-2002
Rattus norvegicus clone CH230-179E15, *** SEQUENCING IN PROGRESS
*** 59 unordered pieces.
ACCESSION AC121387
VERSION AC121387.2 GI:21903324
HTG: HTGS PHASE1.
KEYWORDS
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 193379)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alldrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Blumage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Daylla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunnu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 193379)
Worley,K.C.
Direct Submission
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 193379)
Worley,K.C.
Direct Submission
Submitted (22-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 18, 2002 this sequence version replaced gi:20976316.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWBS
Center clone name: CH230-179E15
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 148042 bases at least Q40
Consensus quality: 148042 bases at least Q30
Consensus quality: 151839 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1183: contig of 1183 bp in length
* 1184
* 1284: gap of unknown length
* 1284
* 2774: contig of 1490 bp in length
* 2774
* 2874: gap of unknown length
* 2874
* 4115: contig of 1242 bp in length
* 4115
* 4215: gap of unknown length
* 4216
* 5507: contig of 1292 bp in length
* 5508
* 5607: gap of unknown length
* 5608
* 7074: contig of 1467 bp in length
* 7075
* 7174: gap of unknown length
* 7175
* 8713: contig of 1539 bp in length
* 8714
* 8813: gap of unknown length
* 8814
* 10113: contig of 1300 bp in length
* 10114
* 10213: gap of unknown length
* 10214
* 11662: contig of 1449 bp in length
* 11663
* 11762: gap of unknown length
* 11763
* 12805: contig of 1043 bp in length
* 12805
* 12905: gap of unknown length
* 12906
* 14173: contig of 1268 bp in length
* 14174
* 14273: gap of unknown length
* 14274
* 15422: contig of 1149 bp in length
* 15423
* 15522: gap of unknown length
* 15523
* 16619: contig of 1097 bp in length
* 16620
* 16719: gap of unknown length
* 16720
* 18244: contig of 1525 bp in length
* 18245
* 18344: gap of unknown length
* 18345
* 19383: contig of 1039 bp in length
* 19384
* 19483: gap of unknown length
* 19484
* 20894: contig of 1411 bp in length
* 20895
* 20994: gap of unknown length
* 20995
* 22250: contig of 1256 bp in length
* 22251
* 22350: gap of unknown length
* 22351
* 24083: contig of 1733 bp in length
* 24084
* 24183: gap of unknown length
* 24184
* 25731: contig of 1448 bp in length
* 25732
* 25731: gap of unknown length
* 28078
* 28178: contig of 2347 bp in length
* 28179
* 29912: gap of unknown length
* 29912
* 30012: contig of 1734 bp in length
* 30013
* 32319: gap of unknown length
* 32320
* 32419: gap of unknown length
* 32420
* 35091: contig of 2672 bp in length
* 35092
* 35191: gap of unknown length
* 35192
* 37133: contig of 1942 bp in length
* 37134
* 37233: gap of unknown length
* 37234
* 38532: contig of 1299 bp in length
* 38533
* 38532: gap of unknown length
* 39980: contig of 1348 bp in length
* 39980
* 40080: gap of unknown length

	*	40081	41886:	contig of 1806 bp in length
	*	41987	41986:	gap of unknown length
	*	41987	44412:	contig of 2426 bp in length
	*	44413	44512:	gap of unknown length
	*	44513	46752:	contig of 2240 bp in length
	*	46753	46852:	gap of unknown length
	*	46853	48812:	contig of 1960 bp in length
	*	48813	48912:	gap of unknown length
	*	48913	51361:	contig of 2449 bp in length
	*	51362	51461:	gap of unknown length
	*	51462	53203:	contig of 1742 bp in length
	*	53204	53303:	gap of unknown length
	*	53304	56572:	contig of 3669 bp in length
	*	56573	56672:	gap of unknown length
	*	56673	59450:	contig of 2778 bp in length
	*	59451	59550:	gap of unknown length
	*	59551	62397:	contig of 2847 bp in length
	*	62398	62497:	gap of unknown length
	*	62498	64638:	contig of 2141 bp in length
	*	64639	64738:	gap of unknown length
	*	64739	67428:	contig of 2690 bp in length
	*	67429	67528:	gap of unknown length
	*	67529	69837:	contig of 2309 bp in length
	*	69838	69937:	gap of unknown length
	*	69938	73355:	contig of 3418 bp in length
	*	73356	73455:	gap of unknown length
	*	73456	77984:	contig of 4529 bp in length
	*	77985	78084:	gap of unknown length
	*	78085	82304:	contig of 4220 bp in length
	*	82305	82404:	gap of unknown length
	*	82405	87045:	contig of 4641 bp in length
	*	87046	91035:	gap of unknown length
	*	87146	91035:	contig of 3890 bp in length
	*	91036	91135:	gap of unknown length
	*	91136	94926:	contig of 3791 bp in length
	*	94927	95026:	gap of unknown length
	*	95027	100061:	contig of 5035 bp in length
	*	100062	100161:	gap of unknown length
	*	100162	104552:	contig of 4291 bp in length
	*	104553	104552:	gap of unknown length
	*	104553	109278:	contig of 4726 bp in length
	*	109279	109378:	gap of unknown length
	*	109379	114411:	contig of 5033 bp in length
	*	114412	114511:	gap of unknown length
	*	114512	119170:	contig of 4659 bp in length
	*	119171	119270:	gap of unknown length
	*	119271	124025:	contig of 4755 bp in length
	*	124026	124125:	gap of unknown length
	*	124126	130339:	contig of 6214 bp in length
	*	130340	130439:	gap of unknown length
	*	130440	135238:	contig of 4799 bp in length
	*	135239	135338:	gap of unknown length
	*	135339	140378:	contig of 5040 bp in length
	*	140379	140478:	gap of unknown length
	*	140479	146853:	contig of 6375 bp in length
	*	146854	146853:	gap of unknown length
	*	146954	152538:	contig of 5585 bp in length

Query Match 7.4%: Score 56.4; DB 2; Length 193379;

Best Local Similarity 44.9%; Pred. No. 0.83;

Matches 213; Conservative 0; Mismatches 261; Indels 0; Gaps 0

OY 186 TTTTGTGTTTGGAACCTTTTGTGTTTTTTCAGCTTTCACAGTAATGCATAA 245
Db 180570 TTGTTATTGTCCTGTGCTGTTGTTATTGTTGCGTTCTGTGTTATTGTTGGCTGTG 180511

OY 246 ATGTCGCCGTTTGGATGAGCTGCTGCAAAAACCTTTTGCTTATGAACTACACAACGACG 305
Db 180510 CTGTCGCTGTTGCTGCTGCTGTTGCTGTTGCTGTTATTTGCTGCTGCTGCTGCTG 180451

OY 306 ATCCATATGACCATTATTTATGAACTTTTATGAGGTTTGGAAATAGCCTCCTTTGCTC 365
Db 180450 TTGTTGTTGTCCTGTGCTGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGCTATGTTGTTGA 180391

QY	366	AAGGAGATATTCCTGGAGCTCTCTATTCCTGGATTTGATGCGGTTGGTATGAGGCTTA	425
DB	180330	TTGTGTCGTGTCGTCTGCTGTTGCTGTGCTGTCGTTGCTGTCGTTGCTGTTGCTGTTG	180331
QY	426	TACTTGCAGCGGGGCTTATTTGGATATCAAAAGCGCTTGATGATTAATCAAAAAAGCTGCTT	485
DB	180330	TTGTGTCGTGTCGTCTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTG	180271
QY	486	TTCAATGCACTTGGGGTAAAGGAGTTATGTAAGCAGGTGTGTTACTATGTCCTGTGACA	545
DB	180270	TTGCTGTTGTTGTTGCTCTGTGCTGTTGCTGTTGTTGCTGTTGCTGTTGCTGTTGCTG	180211
QY	546	GATTAACAGAAATATTTCTCCATTTACATTTGCTAAATAGTTAATAATAGGAAGCTAAAAA	605
DB	180210	TTATTCGCTGTTGCTCTTATTCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	180151
QY	606	ATAGCCTTAATGCTTAACTTGAAGAGATTTGAACCTAGTTTGTGATGCTGCAATGC	659
DB	180150	CTGCTGTTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	180097
RESULT	15		
AC116956/c			
LOCUS	AC116956	133501 bp	DNA
DEFINITION	Dictyostelium discoideum chromosome 2 map 1512060-1645559 strain AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.		
ACCESSION	AC116956		
VERSION	AC116956.1	GI:19919988	
KEYWORDS	HTG; HTGS_PHASE2.		
SOURCE	Dictyostelium discoideum.		
ORGANISM	Dictyostelium discoideum		
REFERENCE	Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.		
AUTHORS	1 (bases 1 to 133501) Gloeckner, G., Eichinger, L., Safranek, K., Pachebat, J., Dear, P., Lehmann, R., Baumgart, C., Farra, G., April, J. F., Guigo, R., Kumpf, K., Tunggal, B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, R. and Noegel, A. A.		
COMMENT	Sequence and Analysis of Chromosome 2 of Dictyostelium Unpublished The Dictyostelium Genome Sequencing Consortium 2 (bases 1 to 133501) Baumgart, C. Direct Submission Submitted (04-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany CDS predictions from Genaid may contain errors. Further Information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml) Funding Agency : Deutsche Forschungsgemeinschaft (DFG). * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. Location/Qualifiers 1. 133501 /organism="Dictyostelium discoideum" /strain="AX4" /db_xref="taxon:44689" /chromosome="2" /map="1512060-1645559" BASE COUNT 50753 a 15240 c 15399 g 51959 t 150 others ORIGIN		
Query Match	7.4%	Score 56;	DB 2;
Best Local Similarity	53.7%	Pred. No. 1.1;	
Matches 116;	Conservative	0;	Mismatches 100;
		Gaps	0;
7	AAAGCAATTCGAATTCGATTAATTTATGCTCAATTAACAAGACTATTTGGAGAGCAATTT	66	

Df 109157 AAAAAAAAAAAAAAGTTAATATTTTTAAAAAAANATTTAAAAAAAAGAATTAATAAACGATA 109098

Oy 67 TCAGCATAATTTGGAAAAAGTTAAATTTAAATACCTTTAAAAACCTTTTTAAATTC 126
| | | | |
Df 109097 AGGAAAAAATAAAAAATTAATAATTTTAAACAACCTCCAAAAAAGCTTTTAATTTT 109038
| | | | |
Oy 127 ATTAATATGCACATAGTACACTTTTAATTAAGSGCTTTTATGAATAACTTTAAAT 186
| | | | |
Df 109037 ATTATTAANAATCAATGCCAATTCCTTAATCATATGATTTTGATATWTAATTTAATTT 108978
| | | | |
Oy 187 TTTTGTTTGGCAACCTTTGTGTGTTTTCAGACT 222
| | | | |
Df 108977 TATTTTATAATTTTATTTTATTTTATTTTATTTT 108942

Search completed: March 27, 2003, 03:05:17
Job time : 2122.82 secs

XX
WPB: 1999-215027/18

DR WPI; 1999-215027/18.

QY	1	ATTGCTAAAGCAATTTGAAATTTGATTTATTTATGCTCAAAATTCAGAAAGCTCTATTGGGAAG	60
	1		
Db	50	ATTATTAAAGCAATTTGAAATTTGCTAAATTTATGCTCAAAATTCAGAAAGCTCTATTGGGAAG	109
QY	61	CGAATTTCAAGCATAATTTTGAAAAAGTTAAATTTAAATTAACCTTTTAAAAACCTTTT	120
Db	110	CGAATTTCAAGCATAATTTTGAAAAAGTTAAATTTAAATTAACCTTTTAAAAACCTTTT	169
QY	121	AATTTCTTAATATGCTACCATAGTACCAAGTTTAAATTAAGGGGTTTTATGCAATAACT	180
Db	170	AATTTCTTAATATGTTACTATATATACCAAGTTTAAATTAAGAGGTTTTATGCAATAAAT	229
QY	181	TTTAATTTTGTGTTTGGCAACCTTTTGTGTTTTTCTTACCTTTGGCTCAAGCTAATGATTC	240
Db	230	TTTAATTTGTTGTTTGGTACCTTTTGTGTTTTTCTTACCTTTGGCTCAAGCTAATGATTC	289
QY	241	TAAAAATGATGCGCTTTGGGATGAGTCTGGAGAAAACTTTTGGTTATGAATAGCA	300
Db	290	TAAAA--GCGCTTTTAATTTGGAGACGGGAGAAAACTTTTAGCTTATGAACATAGTA	346
QY	301	GCAAGATCTATGTGCACATTTTATTTATGAAACCTTTTATAGGTTTGGATAGGCTCCT	360
Db	347	GAAAGATCTATGTGTGCAATTTTATGAAACCTTTTATAGGTTTGGATAGGCTCCT	406
QY	361	TGCTCAAGAGATATTTCTTGAGAGTTCTCTTATTTCTTGAGTTTGAATGCGGTGATAGG	420
Db	407	TGCTCAAGAGATATTTCTTGAGAGTTTCTTATTTCTTGAGTTTGAATGCGGTGATAGG	466
QY	421	GCTTATATCTTGCGGGGCTTATTGATATCAAAAGCGCTTGATGATTACTTAAAAAGC	480
Db	467	GTTAATATCTTACAGAGACCTTATTATATCAAAAGCTCTTGTATAAAGAAATGCTCAAAAGC	526
QY	481	TGCTTTTCATAGGACTTGGGGTGAAGGAGTATGTTAGCAGAGGTTGCTACTATGCTCT	540
Db	527	CGCTTTTAAGTGGCTTGGGGTGAAGGAAATGATTTGGAGGGTGCAGTACTATGCTCT	586
QY	541	GACAAATTAACAGAAATTTATCTTCCATTTTACATTTGCTTAATAGTATTAATAGAGAGCT	600
Db	587	GACAAATTAACAGAAATTTATTTATCTCGTTTACATTTGCTTAATAGTATTAATAGAGAACT	646
QY	601	AAAAAATAGCCTTAATAGTACCTTTAGAGAGATTTGAACCTAGTTTGTATGTCAAATGGG	660
Db	647	GAAAAATAGCCTTAATATATACCTTTTGGAGGGTTTGAGCCAGTTTGAATTTAATATGCG	706
QY	661	CCAATTCAGTGTCTGGGTTGAACGTGCTTCAAAAAAGCATTAATTAATTTA--TTAT	719
Db	707	CCAAGCTAGAGCTCTTGGGTTTGAACATCTTTCAAAAAAGCATTAATTAATTTAATTTAT	766
QY	720	TACAAAAATGGGTGATTTGCAATCTCTATTTGAAATGGT	758
Db	767	TATTAATAATGATGATAGCAATTTTGTATTTGTGATTTGCT	805
RESULT 3			
AA30100			
ID	AA30100	standard; DNA: 749 BP.	
XX	AA30100:		
AC			
XX	17-JUN-1999	(first entry)	
DT			
XX	Borrelia garinii	IP90 protein encoding cDNA.	
DE			
XX	Borrelia; p13 antigen; spirochaete; vaccine; infection; diagnosis; ss		
OS	Borrelia garinii.		
XX			
XX	Key	Location/Qualifiers	
FT	CDS	192..725	
FT		/+tag- a	
XX			
XX	W09912960-A2.		

XX	PD	18-MAR-1999.
XX	PF	04-SEP-1998; 98W0-IB01424.
XX	PR	16-SEP-1997; 97U5-0059036.
XX	PR	10-SEP-1997; 97DK-0001041.
XX	PA	(SYMB-) SYMBICOM AB.
XX	PI	Bergstroem S;
XX	DR	WPI: 1999-215027/18.
XX	P-PSDB:	AAV04280.
PT	Nucleic acid from Borrelia burgdorferi encoding virulence associated protein p13	
XX	Claim 1; Page 113-114; 118pp; English.	
CC	The present sequence encodes a Borrelia garinii IP90 protein. The present invention describes an isolated nucleic acid (I) that: (1) encodes a polypeptide fragment (II) immunologically reactive with rabbit polyclonal antibody raised against a 13 kDa polypeptide of Borrelia burgdorferi B313, but not reactive with most proteins from other spirochaetes; and/or (II) hybridizes under stringent conditions to specific nucleic acid sequences but not to genomic DNA from most other spirochaetes. (I), (II) and transformed cells are useful in vaccines to protect against infection by B. burgdorferi sensu lato. (I), (II) and antibodies raised against (II) are used to diagnose such infections, in standard immunoassays or amplification/hybridization tests. (I) are also used to produce recombinant (II). The 13 kd protein is involved in virulence and is highly conserved within B. burgdorferi sensu lato, but is absent from other Borrelia species (e.g. those responsible for relapsing fever or yaws borreliosis). It should provide a more specific and sensitive antibody response, and diagnosis, compared with use of whole bacteria as antigen.	
XX	Sequence 749 BP: 234 A: 82 C: 150 G: 283 T: 0 other:	
SO	Query Match Best Local Similarity 73.5%; Score 558.2; DB 20; Length 749; Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;	
OY	1 ATTGTTAAAAGCAATTGAATTCATTAATTTTATAGTCGAACATCAAGAAGCTCTATTGGGAAG 60	
Db	24 ATTTGTTAAAAGCAATTTGCAATTTTAAATTTAAGTTTAAATCAAGAAGCTCTATTGGTAAG 83	
OY	61 CGAATTTCAAGCAATAATTTGAAAAAGCTTAAATTTAATAATCATTTTAAAAACCTTTTATTA 120	
Db	84 CGAATTTTGAGTAGACAATTTGAAAAAGCTTAAATTTAATTAAGTTCAAAAAGCCTTTTTTAA 143	
OY	121 AATTTTCATTAATATGCTACCATAGTACCAAGTTTAAATAAAGGGGTTTTTATGAATTAAC 180	
Db	144 AATTTTCATTAATATGCTACCATTAATACAG-TTTAATTAAGGGGTTTTTATGAATTAAG 202	
OY	181 TTTATTTTGTGTTGGGCAACCTTTGTGTTTCTAGCTTGCGTGCACAGTAATGATTC 240	
Db	203 TTTAATTTTATTTTGTGTAATCTTTTGTGCTTTTCTAGTTTGTGCTCAAGATGATTC-- 260	
OY	241 TAAAAATGATGCGTTGGGATGAGTGTGCGAGAAAAACTTTTGGTTTATGAACATAGCAA 300	
Db	261 ----TAAAAGCACTTTAATCTGGGAGCGGAGAAAAAATTTTGGTTTATGAACATTAATA 316	
OY	301 GCACATCTCTAATTTGACACATTTTATTTGAACCTTTTTTATAGGCTTTGCAATAGGCTCCT 360	
Db	317 GAATGATCTCTCTGACACATTTTATTTGAACCTTTTTTATAGGCTTGGGATAGGCTTCTT 376	
OY	361 TGCCGTAAGAGATATCTTGGAGGCTTCCTTATCTTGGAATTTGATGCGGTTGGATATAGG 420	
Db	377 TGCTCAAGAGATATCTTGGAGGCTTCCTTATCTTGGAATTTGATGCGGTTGGATATAGG 436	
OY	421 GCTTATACCTTGGCGGGCTTATTTGGATATCAAAAGCGCTTGATGTATTACTAAAAAAGC 480	

KW gene; ds.
 XX Homo sapiens.
 XX MO200200928-A2.
 XX 03-JAN-2002.
 XX 02-JUL-2001; 2001WO-EP07537.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI: 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 XX Claim 1; SEQ ID NO 1451; 32pp + Sequence Listing; German.
 XX
 XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 XX Sequence 7106 BP; 1752 A; 84 C; 1700 G; 3570 T; 0 other:
 SQ
 Query Match 7.28; Score 54.6; DB 24; Length 7106;
 Best Local Similarity 45.3%; Pred. No. 0.048;
 Matches 198; Conservative 0; Mismatches 239; Indels 0; Gaps 0;
 XX
 QY 1 ATGTGTAAGAAATGGAATGATATTTTATGTCGAATGCAAGAGCTCATTTGGAGG 60
 DB 5916 ATATGTAGCAATATATATGTTAGATATATGAAATTTCTATATATATGTAAGG 5975
 QY 61 CGAATTTCAACCAATATTTGAAAAAGTTAAATTAATTAATTAATTAATTAATTAATTA 120
 DB 5976 ATAGTTTATGCAATATATTTTATACGATTTTAGTGTGTAATTAATTAATTAATTAAT 6035
 QY 121 AATTTCATATATGCTACCTACGATGACGATTTAATAAGGGGTTTATGATTAAC 180
 DB 6036 AATTGGAAGATTTGGCTTTAAATTTGTTTATATATATATAGTTTGTATAAAAATTA 6095
 QY 181 TTTAATTTTGTGTCGACCTTTTGTGTTTTCGATTTGCTCAAGCTAATGATTC 240
 DB 6096 TGTATTTATTTTGTGAGGCTTAATTTTATTTTATTTGATGATTAATAATTTATTTT 6155
 QY 241 TAAAAATGTCGTTGGGAGTACGCTGAGAAAACTTTGGTTATGAACATAGCAA 300
 DB 6156 GTTGTAGGTTTGTGAGATTAATAATGATTAAGATGTGGGAGTGTGTGGTAT 6215
 QY 301 GCAAGATCCATTTTACCATTTTATGACCTTTTATAGGTTTGGAAATAGCTCCTT 360
 DB 6216 ATATGAGGTGATGATTTTGTATTTTATGTTGTTTGTGAGGTTTAAAA 6275
 QY 361 TGTCAAGGAGATATTTCTTGAGGTTCTATTTCTTGATTTGATGCGTTGATATAGG 420
 DB 6276 TGTTAGGCTTTTGGGGCTTTTAGGCTATTTTAAGTTAGTTTATGTTTGAAGAAATA 6335
 QY 421 GCTTATCTTGGGGGG 437
 DB 6336 GTATAGGGGAGGGAGG 6352

RESULT 14
 ABL92319
 ID ABL92319 standard; DNA; 73334 BP.
 XX
 XX AC ABL92319;
 XX
 XX DT 01-JUL-2002 (first entry)
 XX
 XX Chemically treated DNA repair gene fragment complementary to#64.
 XX
 XX DNA repair: cytosine methylation; PMS2L1; PMS2L12; PMS2L3;
 KW PMS2; L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPPL1; RFC4;
 KW DDT1L; FANCB; XRC8; ataxia telangiectasia; aging; Bloom's syndrome;
 KW Cockayne syndrome; Nijmegen breakage syndrome; Werner syndrome;
 KW immunodeficiency; trichothiodystrophy; Fanconi's anaemia; solid tumour;
 KW cancer; ds.
 XX
 XX Unidentified.
 XX
 XX MO200181622-A2.
 XX
 XX 01-NOV-2001.
 XX
 XX 06-APR-2001; 2001WO-EP03972.
 XX
 XX 06-APR-2000; 2000DE-1019058.
 XX 07-APR-2000; 2000DE-1019173.
 XX 30-JUN-2000; 2000DE-1032529.
 XX 01-SEP-2000; 2000DE-1043826.
 XX (EPIC-) EPIGENOMICS AG.
 XX
 XX Olek A, Piepenbrock C, Berlin K;
 XX WPI: 2002-034446/04.
 XX
 XX New nucleic acid derived from genes associated with DNA repair, useful
 PT for diagnosis, e.g. of ataxia telangiectasia, by determination of
 PT cytosine methylation -
 XX
 XX Claim 1; SEQ ID NO 128; 25pp + sequence listing; English.
 XX
 XX The invention relates to nucleic acids containing a sequence of at least
 CC 18 nucleotides of chemically treated DNA of genes associated with DNA
 CC repair, and their complements. The invention also relates to nucleic
 CC acids comprising at least 18 base pairs of the chemically pretreated DNA
 CC of genes associated with DNA repair selected from PMS2L1, PMS2L12,
 CC PMS2L2, PMS2L3, PMS2, L4, PMS2L5, PMS2L6, MGMT, MSH2, NUDT1, TDG, INPPL1,
 CC RFC4, DDT1L, FANCB, or XRC8. Nucleic acids of the invention and related
 CC oligomers, are useful for diagnosis of diseases associated with gene
 CC repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
 CC Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome,
 CC immunodeficiency, trichothiodystrophy, Fanconi's anaemia, solid tumours
 CC and cancer, particularly by determining status of cytosine methylation
 CC and/or by detecting single-nucleotide polymorphisms. Determination of
 CC individual methylation patterns may allow development of individualised
 CC therapies. The sequences given in records ABL923192-ABL92335 represent
 CC chemically pre-treated DNA fragments from genes associated with DNA
 CC repair, and their complements.
 CC Note: The sequence data for this patent is not represented in the
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 SQ Sequence 73334 BP; 24605 A; 858 C; 14432 G; 33439 T; 0 other:
 Query Match 7.18; Score 54.2; DB 24; Length 73334;
 Best Local Similarity 51.0%; Pred. No. 0.078;
 Matches 180; Conservative 0; Mismatches 168; Indels 5; Gaps 2;
 XX
 QY 3 TGTAAAAAGATTTGAATTTGATATTTTATGTCGAATGCAAGAGCTCATTTGGAGCG 62
 DB 33503 TTTTGAGTAAATGAAACGAAATATATATTTTATGATTTTATGAGTAAATAT 33562

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 01:16:47 ; Search time 1065.16 Seconds
(without alignments)
11540.380 Million cell updates/sec

Title: US-09-508-487-18

Perfect score: 759

Sequence: 1 atgttaaagaatgaat.....attctgtatgaatg99tg 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estio:*
4: em_estnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	66.4	8.7	1101	17	CNS0039G
c 2	66.4	8.7	1101	17	CNS001T2
c 3	66.6	8.5	576	17	CNS035N7
c 4	60.6	8.0	1101	17	CNS00E07
c 5	60.6	8.0	1101	17	CNS0100X
c 6	60.4	8.0	1101	17	CNS00E07
					AL063921 Drosophila
					AL078714 Drosophila
					AL0228940 Tetradon
					AL069440 Drosophila
					AL098379 Drosophila
					AL069440 Drosophila

c 7	60.2	7.9	1001	17	CNS0155H	AL105023 Drosophila
c 8	58.8	7.7	661	17	CNS020VJ	AL20980V Tetradon
c 9	57.8	7.6	1101	17	CNS003B0	AL064078 Drosophila
c 10	57.6	7.6	1101	17	CNS000D1	AL065414 Drosophila
c 11	56.8	7.5	970	17	CNS0182E	AL108800 Drosophila
c 12	56.4	7.4	1038	17	CNS01TL7	AL166660 Tetradon
c 13	56.2	7.4	928	17	CNS000DKY	AL071865 Drosophila
c 14	55.8	7.4	1201	17	CNS016FX	AL106695 Drosophila
c 15	55.6	7.3	1101	17	CNS0106X	AL098595 Drosophila
c 16	55.6	7.3	524	17	CNS01U90	AL167541 Tetradon
c 17	55.6	7.3	1203	17	CNS015WU	AL106008 Drosophila
c 18	55.4	7.3	1200	17	CNS016CO	AL106578 Drosophila
c 19	55	7.2	865	17	BH180441	BH180441 016_H_20-
c 20	55	7.2	865	17	CNS07MAN	AL1617393 T3 end of
c 21	55	7.2	1167	17	CNS07360	AL1427102 clone BA0
c 22	55	7.2	1225	17	CNS0161D	AL106171 Drosophila
c 23	54.8	7.2	860	12	BF617324	BF617324 HSVMEC01
c 24	54.8	7.2	952	17	CNS016RS	AL107132 Drosophila
c 25	54.6	7.2	1101	17	CNS001FB	AL060732 Drosophila
c 26	54.4	7.2	1101	17	CNS00LO0	AL068607 Drosophila
c 27	54	7.1	416	17	AZ045571	AZ045571 T234237b
c 28	54	7.1	809	17	B10155	B10155 FAF14-Sp6.1
c 29	53.8	7.1	1038	17	CNS0108N	AL098657 Drosophila
c 30	53.8	7.1	1094	17	CNS012F2	AL101513 Drosophila
c 31	53.8	7.1	1101	17	CNS00HX9	AL073856 Drosophila
c 32	53.8	7.1	1204	17	CNS016E2	AL106628 Drosophila
c 33	53.6	7.1	862	17	CNS029ER	AL187164 Tetradon
c 34	53.6	7.1	868	17	CNS010YV	AL099601 Drosophila
c 35	53.6	7.1	963	17	CNS0075X	AL066857 Drosophila
c 36	53.6	7.1	1184	17	CNS04P4P	AL300850 Tetradon
c 37	53.2	7.0	1101	17	CNS00EVL	AL069706 Drosophila
c 38	53	7.0	1101	17	CNS002J2	AL063072 Drosophila
c 39	52.8	7.0	641	10	AV735039	AV735039 AV735039
c 40	52.8	7.0	987	17	CNS014PQ	AL104456 Drosophila
c 41	52.8	7.0	1125	9	AL547503	AL547503 AL547503
c 42	52.6	6.9	996	17	CNS00EFU	AL071063 Drosophila
c 43	52.6	6.9	1201	17	CNS0167M	AL106396 Drosophila
c 44	52.4	6.9	805	17	AB005977	AB005977 Mouse gen
c 45	52.4	6.9	866	17	CNS006MS	AL065764 Drosophila

ALIGNMENTS

RESULT 1
LOCUS CNS0039G/c 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TERT end of BAC #
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL063921.1 GI:4941778
VERSION AL063921
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
Genoscope.
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial

COMMENT

ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TEf3"

BASE COUNT 201 a 64 c 131 g 202 t 503 others

ORIGIN

Query Match 8.7%; Score 66.4; DB 17; Length 1101;

Best Local Similarity 15.9%; Pred. No. 0.0045;

Matches 103; Conservative 288; Mismatches 253; Indels 2; Gaps 1;

3 TGTAAAGAAATGTAATTTATGTCATCAATCAAGACCTCTATGGGACG 62

1096 WCDTWDNDTRKDDMDKMTWKMDRDRNAGDADRMDAGTWTATMMWMMW 1037

63 AATTCAGCAATTAATGAAAAAGTTAATTAATTAATTAATTAATTAATTA 122

1036 WATMDTMDKMMWMTAAKTDTATMTWRAMRADWAGRDGAKGRDRAATDAGGRD 977

123 TTTCATTAATAGTCACATAGTACAGTTTAAATTAAGGGTTTATGAATAACTT 182

976 GGRKKRDKRDKDDDKGKKRAAKAKWTWMDWMDKMDKMDGADRADD 917

183 TAATTTTGTGGCAACCTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGA 242

916 GAGDDDDGKGDADDDTGTGDDDKMDKMDKMDKMDKMDKMDKMDKMDKMD 857

243 AAAATGTCGCTTGGATGAGTCTGG--AGAAAACTTTTGGTTTGAAGCAAGCA 300

856 DMTTMDAADMDMDRMDAMKMDAMKMDAMKMDAMKMDAMKMDAMKMDAMKMD 797

301 GCAAGATCCTATGTACATTTTATGAACCTTTTATGAGGTTTGAATGCTCTT 360

796 DDKRAADDRDAATWTTTTRDTDWMDKMDKMDKMDKMDKMDKMDKMDKMD 737

361 TGTCAAGAGATATTTCTGAGTCTCTATTTCTGATTTGATGAGTATAGG 420

736 GKMKRTKRRKRRDRDDADADTDARDDRDRDRDRDRDRDRDRDRDRDRDR 677

421 GCTTATCTGGGGGCTTATTTGATATCAAGCGCTTGTATTAATAAAGC 480

676 DRTDMMADAAWTTTDTDTDDMDKRRRRKARRRRRTTAAADWMTWKAMDAM 617

481 TCTTTTCAATGAGTCTGGGTAAGGAGTATGTTAGCAGTGTGTTACTATGCTGT 540

616 KTRARWRWMAADTWTDRKADRDWAKARARARARARARARARARARARAR 557

541 GACAGATTAACAGAAATATCTCCATTTTATGATTTGATTAATAAGAGCT 600

556 WAARAWMAWMAWMTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 497

601 AAAAATAGCCTTATGAGTCTTGAAGATTTGAACCTATGTTT 646

456 AAAAATAAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 451

RESULT 2
CNS00L2 1101 bp DNA linear GSS 14-JUN-1999

DEFINITION Drosophila melanogaster genome survey sequence TEf3 end of BAC:
BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL078714
VERSION AL078714.1 GI:5102004

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission

Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage ;

Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Ooeogawa and

Aaron Mammosen in Pieter de Jong's Laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

ECORI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

pl and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR48P19"
/clone_lib="RPCI-98"
/note="end : TEf3"

BASE COUNT 469 a 6 c 69 g 151 t 406 others

ORIGIN

Query Match 8.7%; Score 66.4; DB 17; Length 1101;

Best Local Similarity 22.6%; Pred. No. 0.0045;

Matches 164; Conservative 220; Mismatches 335; Indels 7; Gaps 1;

7 AAAGATTTGAAATGTAATTTATGTCGAATCAAGACCTCTATTTGGACGCAATT 66

383 AA 442

67 TCAAGCAATTAATGAAAAAGTTAATTAATTAATTAATTAATTAATTAATTA 126

443 AATTTTT 502

127 ATTATATGCTACATAGTACAGTCTTAAATTAAGGGCTTTTATGAATAACTTTAAT 186

503 TTTTATTTTADADKMAAMWDTTTTATWTTTATATATTTTATTTTATTTTAT 562

187 TTTGTTTGGCAACCTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 246

563 TKAKTAAKTAAMAAMWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 622

247 TGTGCGTTTGGAGTGTGCTGGAGAAAACCTTTTGTGTTTATGAACCTACAGCA 306

623 TTTTATTTTAAATTTTATTTTAAATTTTATTTTAAATTTTATTTTAAATTT 682

307 TCCATTTGATACATTTTATTAAGACCTTTTATGAGGTTTAAATGAGCTCTTGT 366

683 AAATTTTAAAMWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 735

367 AGAGATATTTCTTGAGGCTCTTATTTCTGATTTGATGCGGTGATAGGCTTAT 426

736 AWAADKDKDRKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGK 795

427 ACTTGCGGGGCTTATTTGATATCAAGCGCTTGATGTTTATTAATAAAGCTGCTT 486

796 KKAATTTTTRKKKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGK 855

QY	487	TCAAATGAGCTTGGGGTAAAGGAGTTATGTTAGCAGTGTGGTTACTATAGCGTGAACAG	546
		::: ::: : ::: ::: : ::: ::: : ::: ::: :	
Db	856	KTKKKKKKKKKKKGGGKKDDAAAKKKKKKTKKKKKGGGKKGGGKKGGKKDDDAANK	915
QY	547	ATTAAACAGAAATTTATTTCTTCATTTGCTAATTAATAATAGAACTAAAAA	606
		::: ::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :	
Db	916	KKKKKKKKKGGGKKKKGGKKGGKKGGKKTKTKTKKKKKKKKDDAAKGGTKKPAADA	975
QY	607	TAGCCTTATATAGCTTTAGGAGAGATTGAACCTGTTGATGTGTCAGATGGGCCAATC	666
		::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :	
Db	976	AAADMTDAATKKKKKKKKDKDKTKTKKKDDAAATAATKKKKKKKKKKTKKKKKKKKKRD	1035
QY	667	CAGTGCCTTCGGGTGTGAACGTGTCTTTAAAAAAAAGCTATTAAATTTATTTATTAACAAA	726
		::: ::: ::: ::: ::: ::: ::: ::: ::: :	
Db	1036	ADKKKKTKGKKDKGGKGGKGGKKKKKTKTKTKKKKKKKKKKKTKTKTADKKKKKKKKAAD	1095
QY	727	ATGGGT 732	
Db	1096	DKKKK 1101	

LOCUS	DEFINITION	576 bp	DNA	linear	GSS 15-MAY-2000
CNS035N7	Tetradon nigroviridis genome survey sequence PUC-Orig of clone				

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
AL228940	AL228940.1	GI:7887933	GSS: genome survey sequence.	
			Tetradodon nigroviridis.	
			Tetradodon nigroviridis	

REFERENCE
AUTHORS
1 (bases 1 to 576)
Roest-Crollius, H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C.,

TITLE	JOURNAL	REFERENCE	AUTHORS
Human gene number estimate provided by genome wide analysis using Tetraodon n-groviridis DNA sequence	Unpublished	2 (bases 1 to 576)	Roest-Crolius,H., Jallion,O., Dasilva,C., Filames,C., Fisher,C., Bonneau I, Billant A, Ovorot E, Savard M, Boret A and

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 DIRECT SUBMISSION

FEATURES	SOURCE
COMMENT	1 576
trns sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .	Location/Qualifiers

BASE COUNT	ORIGIN	62	a	5	c	18	g	444	t	47	others
/clone_11b="G" /note="Genoscope sequence ID : C0AG214BA03SP1-end : PUC-Or1"											

Query Match	8.5%	Score 64.6	DB 17	Length 576
Best Local Similarity	42.5%	Pred. No. 0.012		
Matches 145, Conservative	30	Mismatches 166	Indels 0	Gaps 0

[illegible]

RESULT 4	CNS00E07	LOCUS	CNS00E07	1101 bp	DNA	linear	GSS 04-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR29P01 of Rpcl-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.						

KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of the nucleotide sequence and cloning of a part of a

The BbGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BbDP *Drosophila* melanogaster BAC library was prepared by Kazutoyo Oseegawa and Aatón Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo.

ECORI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://drososf.mol.uci.edu/drosocn/bac.htm>

BASE COUNT	366 a	66 c	104 g	351 t	214 others
------------	-------	------	-------	-------	------------

Query Match	8.08;	Score 60.6;	DB 17;	Length 1101;
Best Local Similarity	39.48;	Pred. No. 0.056;		

QY	4	GTTAAACAAATGCAATGATATTTTGTGGCAATCAAGAGCTCATTTGGAAGCA	63
		: : : : :	
Db	441	GAKAAATAMTTKGAAGATATPAAAAAARATAAATTAATTAATATATATAA	500
QY	64	ATTCACGCAATAAATTGAAAAAGTAAATTTAAATTAACCTTTAAAAACCTTTTAAAT	123
		: : : : : : : : : : : : : : : : : :	
Db	501	AMTTTAAAAAATAATTAATATAATCAAMTTAAATTAATTTATTAATTAAMWATKE	560
QY	124	TTCAATTAATATGCTCCATGACCAAGTTTATTAATAAGGGCTTTTATGAATTAACCTTT	183
		: : : : : : : : : : : : : : :	
Db	561	TTTTTTTAAATATPAAAAAATAATTTATTTTAAATTTTATATTAATTAAGWAAAAAAMW	620
QY	184	AATTTTGTGTTGGCAACCTTTGTGTTTTTCTAGCTTGTGCTCAAGCTAATGATTTCAA	243
		: : : : : : : : : : : : : : : : :	
Db	621	TATAMHTMTAAATTTATAMTTWAAATTTTAAATAMATTTATTTAAWTTAAWTTTAAW	680
QY	244	AAATGTCGCTTGGAGATGAGTCGCGGAGAAAAACTT	280
		: : : : : : : : : : : : : : :	
Db	681	AMTTTAAATATTTAAARWAAATTTWAAAAAAMAAAMTTTW	717

RESULT	5
CNS0100X	
LOCUS	CNS0100X
DEFINITION	1101 bp DNA linear GSS 26-JUL-1999
ACCESSION	CNS0100X
VERSION	Drosophila melanogaster genome survey sequence SP6 end of BAC BAnoN0G04 of DrosBac library from Drosophila melanogaster (fruit fly) genomic survey sequence.
KEYWORDS	AL098379.1 GI:5609990
SOURCE	GSS.
ORGANISM	Drosophila melanogaster. Drosophila melanogaster.

REFERENCE
AUTHORS
TITLE
JOURNAL

Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
1 (bases 1 to 1101)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peiryogeta;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Epiphytoidea; Drosophilidae; Drosophila.

COMMENT

BP 191 91006 Every cedex - FRANCE (E-mail : seqref@genome.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES	
Source	Location/Qualifiers
1. .1101	
/organism="Drosophila melanogaster"	
/db_xref="taxon:7227"	
/clone="BACN03G04"	
/clone_1b="DrosBAC"	
/plasmid="pBelOBAC11"	
/note="end : SP6"	
195 a	108 c 131 g 161 t 506 others

Query Match	8.0%	Score 60.6;	DB 17;	Length 1101;
Best Local Similarity	16.8%;	Pred. No. 0.006;		
Matches	98;	Conservative 246;	Mismatches 237;	Indels 1;
			Gaps	1;

15 TGAATTGATATTTATGTCAAATCAAGAAGCTCTATTGGGAAGCAATTCAGCAA 74

```

db 425 KGAKWTDTFRDDTWKRTDTIWGKADTWMTDTWKTITWITWITTTGGWKKNDTKTKTKWKD 484

```

DQ 75 TAATTGAAAAAGCTAAATTTTAATAACTTAAAACCTTTTAAATTCATTAATAT 134
 ::||::: ::::|:::| : : : : |:::| :| :| :|
Db 485 DDGGTTWTGTAKDKDXTDKTYVMDNMGWMTDKDKTRAKKRRITDTPTKKTTPRGT 544

OY	135	GCTACCATAGTA-CCAGTTTAAATGAAGGGGTTTTATGAAATTAACCTTTATTTGTT	193
Db	545	KKKIDRCMRKTIIDWGMGEMRADRIDGARADBAANAANKDADGKATDIDATKTIDGHTTKA	604
OY	194	TTGGCAACCTTTGTGTTTTTTTACCTTGTCTCAAGCTAATGATTTCTAAAAATGTCGC	253
Db	605	DTTIDDTDADMMDKMDKDDADDTGARRRGGMGKGGMKKTYTKTKKTKTKTGTDTK	664
OY	254	TTTGGGATGAGTGCCTGGGAAAAAAGCTTTGGTTTGAATAACTAGCAAGCAAGATCCTATT	313
Db	665	WGMKDKWIDDKRDEGRGGDKRDKGTGGDGSKAGDKADWAMWRKATFTAAADATTAATA	724
OY	314	GTACCATTTTATGAACCTTTTATTTAGGCTTTGGAATGAGCTGCTCTTGTCTCAAGAGAT	373
Db	725	AKATMAKTKTKTKTTTTTTTTTTTTTTTTTTTTTTTTTAKRADRDODKARODDKOSDW	784
OY	374	ATTCTTGAGGCTTCCTTATTTCTGGATTTGATGATGCGGCTGTATRAGGGCTATATCTTGG	433
Db	785	DDADDDKMGARDRDDTFRKMGWKRAGKKRKKAKKHDTAAWMAADDTTAATAAKTATA	844
OY	434	GGGGTCTATTTTGGATATCAAGACGCTTATGATTTACTAAAAAAGCTCTTTCATG	493
Db	845	KTGKKAATKDKRGG	904
OY	494	ACTTGGGTAAGGAGATTATGTTAGACAGTGTGCTTACTAGTGTGACAAATTAACA	553
Db	905	AGTDAAGDKKADGGGMMKRRKMSKTKTMTGTGDDTTTGGRRNRGRAGAGRRGR	964
OY	554	GAAATTATTTCCATTACATTTGCTAATGATGTTATATATAG	595
Db	965	KGKAGKAGGKWWKAMADABTWTWMDKMKMTWDMGCK	1006

RESULT 6
CNS00E07/c
LOCUS 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR29P01 of RPC1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069440
VERSION AL069440.1 GI:4949583
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster

REFERENCE
1 (bases 1 to 1101)

COMMENTS	JOURNAL	TITLE	AUTHORS
Determination of this BAC-end sequence was carried out as part of a		Genoscope. .	
		Direct Submission	
		Submitted (02-JUN-1999)	
		Genoscope - Centre National de Sequencage :	
		Bp 191 91006 EVRY cedex - FRANCE (E-mail : sequefr@genoscope.cns.fr	
		- Web : www.genoscope.cns.fr)	

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammaster in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2: cn bw sp, the same strain used for the BDGP's p1 and ESN libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES
source      Location/Qualifiers
1. 1101     /organism="Drosophila melanogaster"
           /db_xref="taxon:7227"
           /clone="BACR29P01"
```

source	1. .661
/organism="Tetradon nigroviridis"	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:45:32 ; Search time 1671.5 Seconds
(without alignments)

13040.962 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749

Sequence: 1 tagaatttcacaataa.....attatctgaanaatgggtg 749

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_ro:*
11: gb_ro:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_mu:*
19: em_mu:*
20: em_mu:*
21: em_mu:*
22: em_mu:*
23: em_mu:*
24: em_mu:*
25: em_mu:*
26: em_mu:*
27: em_mu:*
28: em_mu:*
29: em_mu:*
30: em_mu:*
31: em_mu:*
32: em_mu:*
33: em_mu:*
34: em_mu:*
35: em_mu:*
36: em_mu:*
37: em_mu:*
38: em_mu:*
39: em_mu:*
40: em_mu:*
41: em_mu:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	90.5	678	1 AF085741	AF085741 Borrelia
2	578	77.2	14102	1 AE001117	AE001117 Borrelia
3	541.4	77.7	777	1 AF085740	AF085740 Borrelia
4	513.4	68.5	781	1 AF085739	AF085739 Borrelia
5	98.4	13.1	29766	1 AE000786	AE000786 Borrelia
6	91.6	12.2	52971	1 AE001584	AE001584 Borrelia
7	82.2	11.0	27323	1 AE000789	AE000789 Borrelia
8	71.8	9.6	156533	2 AC117070	AC117070 Dictyoste
9	66	8.8	7110	6 AX251243	AX251243 Sequence
10	66	8.8	7110	6 AX251990	AX251990 Sequence
11	66	8.8	7110	6 AX346458	AX346458 Sequence
12	66	8.8	7110	6 AX349019	AX349019 Sequence
13	66	8.8	34688	6 AX458543	AX458543 Sequence
14	63.8	8.5	7441	6 AX348445	AX348445 Sequence
15	63.8	8.5	7479	6 AX339173	AX339173 Sequence
16	63.8	8.5	14095	6 AX345378	AX345378 Sequence
17	63.8	8.5	15421	3 PFCOMP1RA	X95275 P.falciparu
18	63.6	8.5	178273	2 AC005308	AC005308 Plasmodu
19	63.2	8.4	6131	6 AX345793	AX345793 Sequence
20	62.4	8.3	192994	2 AC109431	AC109431 Rattus no
21	62.2	8.3	6391	6 AX333523	AX333523 Sequence
22	62	8.3	133501	2 AC116956	AC116956 Dictyoste
23	61.8	8.3	14023	6 AX347007	AX347007 Sequence
24	60.2	8.0	9707	6 AX346323	AX346323 Sequence
25	60.2	8.0	164399	3 PFMAL3P6	Z98551 Plasmodu
26	60	8.0	10049	1 AF291051	AF291051 Carsonell
27	60	8.0	43993	2 AC116965	AC116965 Dictyoste
28	60	8.0	107289	2 AC116923	AC116923 Dictyoste
29	60	8.0	349980	6 AX344569	AX344569 Sequence
30	59.8	8.0	66910	2 AC112686	AC112686 Mus muscu
31	59.6	8.0	36601	2 AC116981	AC116981 Dictyoste
32	59	7.9	3873	6 AX458573	AX458573 Sequence
33	59	7.9	5675	6 AX346178	AX346178 Sequence
34	59	7.9	6464	6 AX345417	AX345417 Sequence
35	58.8	7.9	1192	9 HSA323759	AJ323759 Homo sapl
36	58.8	7.9	170654	9 AC073901	AC073901 Homo sapl
37	58.6	7.8	1589	3 AF275222	AF275222 Schlechte
38	58.6	7.8	9731	6 AX345892	AX345892 Sequence
39	58.4	7.8	8237	6 AX251560	AX251560 Sequence
40	58	7.7	15387	6 AX345086	AX345086 Sequence
41	57.8	7.7	93475	5 AL672088	AL672088 zebrafish
42	57.8	7.7	349980	6 AX344552	AX344552 Sequence
43	57.6	7.7	7781	6 AX346071	AX346071 Sequence
44	57.4	7.7	604	3 DIN418038	AJ418038 Dycsla In
45	57.2	7.6	13125	6 AX281445	AX281445 Sequence

ALIGNMENTS

RESULT 1
AF085741 678 bp DNA linear BCF 24-APR-2001
DEFINITION Borrelia garinii membrane protein p13 (p13) gene, complete cds.
ACCESSION AF085741
VERSION AF085741.1 GI:4731124
KEYWORDS
SOURCE Borrelia garinii.
ORGANISM Borrelia garinii.
Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.
REFERENCE 1 (bases 1 to 678)
AUTHORS Noppa,L., Ostberg,Y., Lavinovich,M. and Bergstrom,S.
TITLE p13, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains

Pred. No. is the number of results predicted by chance to have a

JOURNAL Infect. Immun. 69 (5), 3323-3334 (2001)
 MEDLINE 21189251
 PUBMED 11292755
 REFERENCE 2 (bases 1 to 678)
 AUTHORS Noppa, L., Ostberg, I., and Bergstrom, S.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea University, Umea S-901 87, Sweden
 FEATURES
 source
 1. 678
 /organism="Borrelia garinii"
 /strain="1p90"
 /db_xref="taxon:29519"
 /chromosome="large linear"
 121. 654
 /gene="p13"
 121. 654
 /gene="p13"
 /codon_start=1
 /transl_table=1
 /product="membrane protein p13"
 /protein_id="AAB28352.1"
 /db_xref="GI:4731125"
 /translation="MKKFLFIVIFCAFSFQODSKSTFENIGAGEKFLVETNKKD
 SLVFLNLFGLGIGSFAGDILIGSLIFGDAVIGLILTGAYIDIKPDNNAKKA
 DFKWTGKGMVLGVVMAVTRLEIVLPPTFANNYNRKLNINIALGFEPSFDIN
 MGQSAAGFSLSRKSY"
 BASE COUNT 202 a 77 c 140 g 259 t
 ORIGIN
 Query Match 90.5%; Score 678; DB 1; Length 678;
 Best Local Similarity 100.0%; Pred. No. 1.4e-101;
 Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 ATGGGCCAAGCTAGTCTCTGTTTGGACTGCTTTCAGAAAAAGCTATTAAATTTAT 731
 Db 601 ATGGGCCAAGCTAGTCTCTGTTTGGACTGCTTTCAGAAAAAGCTATTAAATTTAT 660
 QY 732 TTATCTAGAAAATGGCTG 749
 Db 661 TTATCTAGAAAATGGCTG 678
 RESULT 2
 AE001117/c 14102 bp DNA .linear BCT 15-DEC-1997
 LOCUS
 DEFINITION Borrelia burgdorferi (section 3 of 70) of the complete genome.
 ACCESSION AE001117 AE000783
 VERSION AE001117.1 GI:2687907
 KEYWORDS
 SOURCE
 ORGANISM
 Borrelia burgdorferi.
 Borrelia burgdorferi
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
 burgdorferi group.
 1 (bases 1 to 14102)
 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A.,
 Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,
 Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D., Richardson, D.,
 Peterson, J., Kerlavage, A.R., Quackenbush, J., Salzberg, S.,
 Hanson, M., van Vugt, R., Palmer, N., Adams, M.D., Gocayne, J.,
 Venter, J.C. et al.
 Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi
 Nature 390 (6660), 580-586 (1997)
 9403685
 2 (bases 1 to 14102)
 Fraser, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R.A.,
 Lathigra, R., White, O., Ketchum, K.A., Dodson, R., Hickey, E.K.,
 Gwinn, M., Dougherty, B., Tomb, J.F., Fleischmann, R.D.,
 Richardson, D., Peterson, J., Kerlavage, A.R., Quackenbush, J.,
 Salzberg, S., Hanson, M., van Vugt, R., Palmer, N., Adams, M.D.,
 Gocayne, J., Weidman, J., Uterback, T., Wathey, L., McDonald, L.,
 Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M.D., Horst, K.,
 Roberts, K., Hatch, B., Smith, H.O. and Venter, J.C.
 Direct Submission
 Submitted (12-DEC-1997) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 TITLE
 JOURNAL
 FEATURES
 source
 1. 14102
 /organism="Borrelia burgdorferi"
 /db_xref="taxon:139"
 95. 733
 /gene="BB0027"
 95. 733
 /gene="BB0027"
 /note="hypothetical protein; identified by glimmer;
 putative
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB0027"
 /protein_id="AAC66429.1"
 /db_xref="GI:2687920"
 /translation="MKRYIFILAVLLIGVNIKIAANIDPHTNSTIGIDISVGI
 PFIYNDLSKATPIPLVFGGIGALIKYQHILNNLIGLEKRMFEDINHSFNILNPSS
 SVGRIFYSPVTFPSINIFIDIGELFQIPVFNIGFSLNTYGDNRNNNTLNTEDALPT
 ISFGSGLMNNENYMARCATASMMWMEFEGNSAKMAHFAVLSLVTVNVKL"
 741. 1790
 /gene="BB0028"
 741. 1790
 /gene="BB0028"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=1
 /product="B. burgdorferi predicted coding region BB0028"

/protein_id="AAC66428..1"
 /db_xref="GI:2687919"
 /translation="MKQRYENFKRRLILNDLIFLLACSSSESIFSQGNLOKIKHEYY
 NLLGSSSPGSLVGETLTYIAAMHLFKKENKIKIEIDLSNSEYFINDIVNSGTYLL
 ANKNEEELPSEKNDWTLKFKRLKAYRFLKSGRGVGEAYILADKNNREKIFD
 LOGSKTPQATQENDKRFQIJSNEELTIGNSLKITQMOMNNNTYTIIDYOQAKELMP
 IITTSIRSSSEYLVWGTGYNLDDTRFKYSTNTNNTTPIPIODEYGESSSTFAEEDDALL
 ICSNNGFAEFTKNKEGIFALRAPSKSVPEGAVNGSQSLGTANDIIPVSNTYIILTO
 GGLKMLKRNKRLTKE"
 complement(1776..2207)
 /gene="BB00429"
 complement(11776..2207)
 /gene="BB00429"
 /note="similar to SP:p17868 GB:X03897 PID:40141
 PID:1303831 GB:AL009126 percent identity: 32.58;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAC66423.1"
 /db_xref="GI:2687914"
 /translation="MLSKIFWDADSCNFKIIFLOKFLHNKSKLIYVSNKFELETT
 ENVALEVDNDQSFILNLDARSIVTIDEDIFAKLLDGLGQVANDSGRIFDININNY
 LYFRSKINFLNLIKIKTKKYDDEFNKSRKEKFTNPFSLFS"
 complement(2197..2832)
 /gene="BB00350"
 complement(2197..2832)
 /gene="BB00350"
 /note="similar to GB:Z36269 PID:1049113 percent identity:
 30.95; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="signal peptidase I (lepb-1)"
 /protein_id="AAC66422.1"
 /db_xref="GI:2687913"
 /translation="MAYPLTEBQRLRKQRQKVFKKFLFLLINFEFFKRYLOIEMI
 KSNEMLPITKNASLEFVATHITSEFPLKANDVLVEDFLSNPFLLTLIKDFEFLN
 KLFKRSKRSKRIAAVQDGSYVRGLNVLVNRKDTNFEYLNGLNYSYKLLDGFENDEE
 VIKCFCTLKNEFFLLNDNISLVNDSRIRGPIKNNAIVSFLVAKVVDKYIK"
 complement(2834..3814)
 /gene="BB0031"
 complement(2834..3814)
 /gene="BB0031"
 /note="similar to PID:1652260 percent identity: 35.98;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="signal peptidase I (lepb-2)"
 /protein_id="AAC66421.1"
 /db_xref="GI:2687912"
 /translation="MYLRLDKPASFVLIHVEIKLYTRKKRYCKRLAKRKGRLNF
 LEDFVAALIFVILYNQFVOYAKIIDGSEMENTLOIGPFLPDKRSYGPPELLPGFLKN
 GKRAEBSDDITIFENPEYKSKGVFFDIFORILMLITLSFDLDDDEGVGNPRVRLVLR
 GAFAGKATYFRENKRAYIKRGEENFLIEDSYWOLVKNFKIRKIVANEDYGIYGDGA
 MGVALLSQNLINLSITPDSFEDVARMIDFEERERLEYKLSAFMRYVANEKALIMRDY
 GIYVVDGYILPIGDNRDNSHDGRFEGYVINKNKVGLRLLIYLPFSRVGFI"
 3885..5357
 /gene="BB0032"
 3885..5357
 /gene="BB0032"
 /note="hypothetical protein; identified by glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB0032"
 /protein_id="AAC66427.1"
 /db_xref="GI:2687918"
 /translation="MKEVDENSANLEFEVKLPILGIEPKVYVFLTIILLSLISFL
 IITPEKRGVAYLKINSIENTYIYLNKRYIGRPILNKYINATGAGYARAKRMGKTYE
 ORIKHNKFNQYSLOINLELYDPEKILIKOKRSLIMYKIKNTENYKILIPFSLS
 SELKHPYIKKPLKDSIPYLNSTDMRFDPLNSKATISIDQDNSNDBELWNSLAKTFE
 DLNNAITQVNELENDKDLKILTKNAPVWTKLAKLDENQIOLSKNKRINIKLGEFK
 INSNATIEQVNELENDKSNISLSTYNNVEFLQDQNTWYKEYODFLKENPFWALN

gene
 CDS
 5369. 5821
 /gene="BB0033"
 /note="similar to GP:2073503 percent identity: 47.68;
 identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="small protein (smpB)"
 /protein_id="AAC66420.1"
 /db_xref="GI:2687911"
 /translation="MTNTNLLLENKKRAKNFYIEKISCGIVLKGTEVSKKKKLSF
 NNSVASIKKEELMLENHLVSKYKGNIFNDELBPRLILKKRELQALCKFKKEGGT
 LIPISFLKSIIVKEVIGICGKKLYRKRELKKOKSIKKDLSREIKTK"
 complement(5892. 6431)
 /gene="BB0034"
 complement(5892. 6431)
 /gene="BB0034"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB0034"
 /protein_id="AAC66426.1"
 /db_xref="GI:2687917"
 /translation="MKNKLIFFVLTFCVFFSSFAQANSKGARGMSGKELVYETSK
 QDPVLPFLNLFEGEGIGSFAGDIIIGSLIGDVGIGILAGVLDIKALDGIYK
 KAAQWMTGKGVMLAGVMTAVVRLTEIILPFTFANSYNRKLKNSLVAGGEPSPD
 VAMQSSALGFELSFKKSY"
 complement(6499. 8379)
 /gene="BB0035"
 complement(6499. 8379)
 /gene="BB0035"
 /note="similar to GP:1405462 percent identity: 31.66;
 identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="DNA topoisomerase IV (parC)"
 /protein_id="AAC66419.1"
 /db_xref="GI:2687910"
 /translation="MDIRTVLKNDFLOYSSVYIKRAIASVVDGKFPVORRIHSLEP
 MHQGNPRKNAVYNGMTKKYHPHGDSTYEALVNIANNDDPLEQNGNGNLPFGDPRASA
 SRIECHKPLAPLVYLSKETIETESTYDGRNNEPLLPAKIPVILIQSSGELAVGWA
 AKILPHNENFLNVAKSELLGESYDIYDPEPTGGIYVNVNADONGVLAIVRAIETID
 EKIIVIELRELPGGETESLISIEKAIRKNYIKVSIINDFTAEVNAIELSLPRGVYASE
 VIEKLYVNTCOISISVNLILLSERYPVVYTIKDLIKFHAHLQKLIKMELEOKSKI
 LKIPFETVLEQIFIEKKIYKLEIETISKEENLISIISEVLRHKSFESEVLEKEDVET
 LKIPKIKSLIEDIDKNSKDKIINKELKSIINSNTSGYSINPIDILLAKYSEKHOR
 KTKISILSKKVKETATRNKKVYINLKEGFRGTSLPFGETIGNASTYDKILVRENSY
 VLNIEKRTYIDKRNVCALYDINNKSQEQFISIIYFRLNPFYVVKFKDKIDKTYTDV
 YFKEGNDERFVDSLNPEVFEFTSNKDIVRIEIDNEMVKSRSISNNLKKVK
 FK"
 complement(8379. 10178)
 /gene="BB0036"
 complement(8379. 10178)
 /gene="BB0036"
 /note="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018
 PID:467396 percent identity: 37.98; identified by sequence
 similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="DNA topoisomerase IV (parE)"
 /protein_id="AAC66418.1"
 /gene="BB0037"
 complement(8379. 10178)
 /gene="BB0037"
 /note="similar to GB:D26185 SP:P05652 GB:X02369 PID:40018
 PID:467396 percent identity: 37.98; identified by sequence
 similarity: putative"
 /codon_start=1
 /transl_table=11
 /product="DNA topoisomerase IV (parE)"
 /protein_id="AAC66418.1"

Db	6623	TAGAAATTTTCACAAATTAAGACATGTTTAAAGAAATTGAATTCATATATTTTAAGTCA	6564
Qy	61	AATCAAGAAGCTCTATTTGCTAAGCAATTTGCAATTAATTTGAAAAAGTTAAATTTA	120
Db	6563	AATCAAGAAGCTCTATTTGCTAAGCAATTTGCAATTAATTTGAAAAAGTTAAATTTA	6504
Qy	121	AATATTTCCAAAGAGCTTTTAAATTTCAATTAATATGCTACCAATATACCAAGTTTAT	179
Db	6503	AATATCTTTAAAAACCTTTTAAATTTCAATTAATATGCTACCAATATACCAAGTTTAT	6444
Qy	180	AAAGGAGTTTATTAAGATTAAGTTTAAATTTTAAATTTTGGTAATCTTTTGCCTTTCT	239
Db	6443	AAAGGAGTTTATTAAGATTAAGTTTAAATTTTAAATTTTGGTAATCTTTTGCCTTTCT	6384
Qy	240	AGTTTGGCTCAAGATGATCT-----AAAGACATTTTAATCTGGGAGCGGGAGAAAA	293
Db	6383	AGCTTTGGCTCAAGATGATCT-----AAAGACATTTTAATCTGGGAGCGGGAGAAAA	6324
Qy	294	TTTTTGGTTTGAAGCTAATTAAGAAAGATTCCTCTGTACCAATTTTATTTGAACCTTTT	353
Db	6323	CTTTTGGTTTGAAGCTAATTAAGCAAGCAAGATTCCTGTACCAATTTTATTTGAACCTTTT	6264
Qy	354	TTAGGTTTCGGGATAGATGATCTTTTGCTCAAGAGATATCTTTGAGAGTCTCTTATTCCT	413
Db	6263	TTAGGTTTCGGGATAGATGATGCTCTTTTGCTCAAGAGATATCTTTGAGAGTCTCTTATTCCT	6204
Qy	414	GGATTTGATGCGGTTGCTGTATAGGTTTAACTTACACAGACCTTATTTGGACATCAAGAT	473
Db	6203	GGATTTGATGCGGTTGCTGTATAGGCTTTTACTTGGGCGGCTTATTTGGATATCAAAAGC	6144
Qy	474	TTTGATTAATTAAGCTCAAAAGACCTATTTTAAGTGCACCTGGGGTAAAGGAATGATGTTG	533
Db	6143	CTTGATGATGATTAACATAAAAGCTCTCTTTTCATGAGACTTTGGGGTAAAGGAATGATGTTG	6084
Qy	534	GCAGGTGTTGTTACTATGAGCTGTGACAAAGATTGACAGAAATGTTCTTCATTTACATTT	593
Db	6083	GCAGGTGTTGTTACTATGAGCTGTGACAAAGATTGACAGAAATGTTCTTCATTTACATTT	6024
Qy	594	GCTATTAATTTAATACAGAAAGCTGAAAAATAGCTTAATATATGCTTTGGAGAGATTGAG	653
Db	6023	GCTAATTAATTTAATACAGAAAGCTGAAAAATAGCTTAATATATGCTTTGGAGAGATTGAA	5964
Qy	654	CCTAATTTTGAATTAACATGGGCCAAGCTAGTGCCTCTGTTGGATGCTCTTCAAA	713
Db	5963	CCTAATTTTGAATTAACATGGGCCAATTCAGATGCTCTCTGTTGGATGCTCTTCAAA	5904
Qy	714	AAAAGCTAATTAATTTTATCTAGAAAAATGGGTG	749
Db	5903	AAAAGCTAATTAATTTTATCTAGAAAAATGGGTG	5868
RESULT 3			
LOCUS	AF085740	777 bp	DNA linear BCT 24-APR-2001
DEFINITION	<i>Borrelia afzelii</i> membrane protein p13 (p13) gene, complete cds.		
ACCESSION	AF085740		
VERSION	AF085740.1	GI:4731122	
KEYWORDS			
SOURCE			
ORGANISM	<i>Borrelia afzelii</i> . <i>Borrelia afzelii</i> . Bacteria; Spirochaetales; Spirochaetaceae; <i>Borrelia</i> ; <i>Borrelia</i> <i>burgdorferi</i> group. 1 (bases 1 to 777) Noppa, L., Ostberg, Y., Lavinovich, M. and Bergstrom, S. p13, an integral membrane protein of <i>Borrelia burgdorferi</i> , is C-terminally processed and contains surface-exposed domains <i>Infect. Immun.</i> 69 (5), 3323-3334 (2001)		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
Submitted	(24-AUG-1998)	Microbiology, Medical Microbiology, Umea	

FEATURES		University, Umea S-901 87, Sweden
SOURCE		Location/Qualifiers
		1..777
		/organism="Borrelia afzelii"
		/strain="ACAI"
		/db_xref="taxon:29518"
		/chromosome="large linear"
		121..657
gene		/gene="p13"
		121..657
CDS		/gene="p13"
		/codon_start=1
		/transl_table=1
		/product="membrane protein p13"
		/protein_id="AAD28361.1"
		/db_xref="GI:4731123"
		/translation="MNKELIVLAFVCFSSFAQADSKAFNLGAEKLLAETSKK DPVPEFLNFIIGIGSFAGQDILGFFLLGEPAVIGILITGAYLIDIALKNAPK AAFWKMGKMLAGAVTMAVTRLEIILIPTFEANSYNRLKNSLNIAFGFERSFDI NMGQASLGFELSPKRSY"
BASE COUNT	228 a 84 c 148 g 316 t	1 others
ORIGIN		
Query Match	72.3%; Score 541.4; DB 1; Length 777;	
Best Local Similarity	89.6%; Pred. No. 2.9e-79;	
Matches 605; Conservative 0; Mismatches 66; Indels 4; Gaps 2;		
QY	73	CTATTGGTAAGCGAATTTGAGTAACAATTTGAAAAAGTTAAATTAATAGTCCAA 132
Db	1	CTATTGGGAGCGAATTTCAAGTAACTACTTTGAAAAAGTTAAATTAATAGTTTAAA 60
QY	133	AGCCTTTTAAATTCATTATATATGACCAATATACAG-TTAAATAAAGGGTTTTT 191
Db	61	AACCTTTTAAATTCATTATATGATGACCTATATACAGTTTAATTAAGAGGTTTTT 120
QY	192	ATGAATTAAGTTTATTTTATTTTATTTGGTAATCTTTTGCTTTTCTAGTTTGGTCAA 251
Db	121	ATGATTAATTTTATTTATTTTATTTTGGTATCTTTTGCTTTTCTAGTTTGGTCAA 180
QY	252	---GATGATCTTAAAGACACTTTTATATCTGGAGCGGAGAAAAATTTGGTTATGAA 308
Db	181	GCTGATGATCTTAAAGAGCGCTTTTATTTGGAGCGGAGAAAACTTTTATGCTTATGAA 240
QY	309	ACTAATTAAGAAAGATTCCTCTGACATTTTATTTGAACCTTTTATTTGGGTTGGGATA 368
Db	241	ACTAGTAAGAAAGATTCCTCTGACATTTTATTTGAACCTTTTATTTGGGTTGGGATA 300
QY	369	GTTCTTTTGTGCAAGAGATATCCTTGAGAGTCTCTTATTTCTTGATTTGATCGGTT 428
Db	301	GTTCTTTTGTGCAAGAGATATCCTTGAGAGTCTCTTATTTCTTGATTTGATCGGTT 360
QY	429	GGTATTAAGGTTAATCTTACAGAGCTTATTTTGGACATCAAGATTTTGAATTAATAGCT 488
Db	361	GGTATTAAGGTTAATCTTACAGAGCTTATTTTGAATCAAGCTTGAATTAAGATGCT 420
QY	488	AAAAAAGCTGATTTTAATGAGCACTGGGGTAAGGAAATGATGTGGCAGGTGCTACT 548
Db	421	CCAAAGAGCGCTTTTAAAGTGCATGGGGTAAGGAAATGATGTGGCAGGTGCTACT 480
QY	549	ATGGGTGTGACAAAGTTGACAGAAATTTGTCATTTCAATTTGAATTAATTAATAC 608
Db	481	ATGGGTGTGACAAAGTTGACAGAAATTTTATTTCCGTTTACCTTTGCTAATTAATTAAT 540
QY	609	AGGAAGCTGAAAAAATAGTCTTAATTAATAGCCTTTGGAGATTTGAGCCTTAGTTGATTT 668
Db	541	AGGAAGCTGAAAAAATAGCCTTAATTAATAGCTTTTGGAGGTTTGAGCCAGTTTGAATTT 600
QY	669	AACATGGGCCAAGCTAGTGCCTTGCTGGTTTGGAGCTGCTTTCAAAAAAAGCTAATTAATTT 728
Db	601	AATATGGGCCAAGCTAGGCTCTTGCTGGTTTGAACATCTTTCAAAAAAAGTTAATTAATTT 660
QY	729	TATTTATCTAGAAA 743

	DB	661	TATTTATTATTAA	675
	RESULT 4			
	AF085739		781 bp	DNA linear BCT 24-APR-2001
	LOCUS	Borrelia burgdorferi membrane protein Pl3 (pl3) gene, complete cds.		
	DEFINITION			
	ACCESSION	AF085739		
	VERSION	AF085739.1	GI:4731120	
	KEYWORDS			
	SOURCE			
	ORGANISM	Borrelia burgdorferi. Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group. 1 (bases 1 to 781) Noppa,L., Ostberg,Y., Lavrinovich,M. and Bergstrom,S. Pl3, an integral membrane protein of Borrelia burgdorferi, is C-terminally processed and contains surface-exposed domains Infect Immun. 69 (5), 3323-3334 (2001)		
	JOURNAL	21189251		
	MEDLINE	11292755		
	PUBMED	2 (bases 1 to 781)		
	REFERENCE	Noppa,L., Ostberg,Y. and Bergstrom,S. Direct Submission Submitted (24-AUG-1998) Microbiology, Medical Microbiology, Umea University, Umea S-901 87, Sweden Location/Qualifiers 1..781 /organism="Borrelia burgdorferi" /strain="B31" /db_xref="taxon:139" /chromosome="large linear" 122..661 /gene="pl3" 122..661 /gene="pl3" /note="similar to Borrelia burgdorferi predicted coding region BB0034 encoded by Genbank Accession Number AE001117"		
	CDS	/codon_start=1 /product="membrane protein pl3" /protein_id="RAD28360.1" /db_xref="GI:4731121" /translation="MKRLIFVLATPCVFSSQANDSKNGAFMSAGELLYETSK ODPFLPLNFLFGISGFQGDILGSLIGDAVIGIILAGAVLDIKALDITRK KAAOFTWKGVMYAGVTMAVRMLTELILPFPTANSYNKLKNSLVNALGGPEPSFD VAMQSSALGPSEFKSY"		
	BASE COUNT	223 a 98 c 153 g 307 t		
	GIN			
	Query Match	68.5%; Score 513.4; DB 1; Length 781; Best Local Similarity 86.4%; Pred. No. 1.1e-74;		
	Matches	592; Conservative 0; Mismatches 86; Indels 7; Gaps 2;		
OY	72	TCTATTGGTAAGCGAATTCGAGTACAATTTGAAAAAGTTAAATTAATAGTTCCAA	131	
Db	1	TCTATTGGGAAGCGAATTCGAAGCAATAATTTGAAAAAGTTAAATTAATTAACTTTAA	60	
OY	132	AAGCCTTTTTRAAATTTATTAATATGCTACCATAATACCAAG-TTTAAFAAGGGGT	190	
Db	61	AAACCTTTTTTAAATTTCTATTAAATATGCTACCATGATACAGTTTAAFAAAGGGGT	120	
OY	191	TATGAATTAATTTTAAATTTTATTTTATTTTGATATCTTTGTCTTAGTTGCACA	250	
Db	121	TATGAATTAATTTTAAATTTTGTGTTGGCAACCTTTTGTGTTTTTCTAGCTTGTCCA	180	
OY	251	AGATGATTC-----AAAAGACATTTTAATCTGGGAGCGGAGAAAAATTTTGTTA	304	
Db	181	AGCTAATGATTTCTAAAAATGCGCTTTGGGATGAGTCTGAGAGAAAACTTTTGTTA	240	
OY	305	TGAACATAATAAGAAGATTCCTTGTCACATTTTATTTAGAACCTTTTATAGGGTGG	364	

LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
LOCUS AE000786	Borrelia burgdorferi plasmid lp28-2, complete plasmid sequence.	AE000786			
VERSION AE000786.1	GI:2690008				
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					
gene					

CDS
116..1009
/gene="BBG01"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG01"
/protein_id="AAC6073.1"
/db_xref="GI:2690030"
/translation="MRKSLFLYLTLGLMGSCNLSKLSNKEOKNNNNVKEVNSVQ
EDGLNDLYSNQKSEFTKNQKMGKEDLINPEPIIPESPKANIPINISIVHOK
KEIKEDDLIPSTNEKEADEAIKYLEENTLONSPSELIRREVVLKDEVALISDYD
VIEKINRKTSLSENNKNNRDKINKTITLQNNIKINIELFQILNMDIAENETRSAR
FEDTAQRLKESLIRKLESNNRSTYALSLROALSDANSALSLESFAKRAEPNVR
KKRIKELIRHAKVLESLNKK"
1047..1928
/gene="BBG02"
1047..1928
/gene="BBG02"
/note="similar to PIR:A64689 percent identity: 35.78;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66053.1"
/db_xref="GI:2690010"
/translation="MEINLOSKLNKNNKLIFFISCSLVSTREPDRNFETYSKNR
GVIRPGYKIMKHILEIONNIALITRLSKTDFRSHAFVSKISELIPLGVFEPY
IODESEPERVKENKFRHLVDYKNNKFTAEELIGYVAIILSNIRDEYELI
QIDEPKIEVDNSDIPYTLGKIGTDLINSLVAVPININIKCSFDELTKONSI
IEKVFYKRENNELYYNQTSPFINVSKENVYVITGSMTLSKYLTRKGRMPKREVEH
LEKVIKIHITIGIOKK"
2104..2495
/gene="BBG03"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact.; similar to
GP:2239281 percent identity: 63.64; identified by sequence
similarity; putative; conserved hypothetical protein,
authentic frameshift"
complement(2750..2857)
/gene="BBG04"
complement(2750..2857)
/gene="BBG04"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG04"
/protein_id="AAC66074.1"
/db_xref="GI:2690031"
/translation="MKQEAISITFDLEIAVHIKTLIDINLILYINGV"
complement(2891..4056)
/gene="BBG05"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact.; similar to
GP:228714 percent identity: 99.12; identified by sequence
similarity; putative; transposase-like protein, authentic
frameshift"
4208..5368
/gene="BBG06"
4208..5368
/gene="BBG06"
/note="similar to GB:U03641 PID:458216 percent identity:
66.76; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66054.1"
/db_xref="GI:2690011"
/translation="MKVFFPLKLCITINPINIRSSMTMNNKKPNCHNLOOKLIY
LLSLVAVNSKYNKYTOKNILYFENENKRGQDFTLRMONTLYLREKYEKATYV
YKHGINNGETIYKLNKSKKECYIKNQHREKKDYFRKRVNLYNDKNNKNSVD
LVLCNNNNIKERKTIQIEKYOVAKYFNKCNPSLKEITPLINDINDKDELKIL

gene
CDS
5378..5950
/gene="BBG07"
5378..5950
/gene="BBG07"
/note="similar to PID:1143398 PID:1174342 PID:1174350
PID:2182755 percent identity: 57.84; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66055.1"
/db_xref="GI:2690012"
/translation="KQNNKASTQLVKPIYRCSNKKDLFIKIKONDKTYHKKIMDI
YKGLNKKNNKYRISLRLGLENOSKVEENLFTLRADKDLGITGKKRIKILPYQY
VGGIEKYSLSKTYMEVREKRSIPECYRSLIRLKENVNTYKALFSMETTLER
OYVDFKKYVPOKGPFIKMEKSMKNHIL"
5911..6678
/gene="BBG08"
5911..6678
/gene="BBG08"
/note="similar to PID:1174351 percent identity: 39.26;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="stage 0 sporulation protein J (spooJ)"
/protein_id="AAC66056.1"
/db_xref="GI:2690013"
/translation="MDRKLAKKPHITAIASIGVGKSTSSIMESTILSKNNVLLY
DLPONAVNSYFTTODHPRELININYSLSIKHKTFKQVVISISKNDFIPSYEL
AKRSKGNOKKELMIRNAVNYLEDYDVIDTPDPSSELNALYADKVIIPVLE
RMAVENPLINQIKELENNFGKREKAIHIRSKYDIGNVSTELMSLKEVYLNKF
IGEVHSEALKRIIDVAIGPENENYKEYLILEKI"
6737..7288
/gene="BBG09"
6737..7288
/gene="BBG09"
/note="similar to PID:1143400 percent identity: 43.68;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC66057.1"
/db_xref="GI:2690014"
/translation="MVNKRKYIINDIRVNAVYINTEERDKREVELKLNRIEDD
IRNKNITMKLILEIRNKLILDGKFFEDFIDFKIARQAVKIKIADLIFEGKLE
EDIIENGIDKTLFNLMDKKINSKANLIPPLVRLETOGACDFYKMPKFANYILED
FYCKNRQLIKRILEYKNNKOKYS"
complement(7483..10779)
/gene="BBG10"
complement(7483..10779)
/gene="BBG10"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=11
/product="B. burgdorferi predicted coding region BBG10"
/protein_id="AAC66075.1"
/db_xref="GI:2690032"
/translation="MKLDELITPLISMSINNOUKDAIETLTKRIABQFSSLDLKS
LEKSAKGSNIEKRLNKSNAIKNPFSLANSVNSVSKAGSMKSGIKVLKSGKGG
SVKNLANKTGAEFQOMLAARPIIMAKVIAQIGSTISGIFDGAMDLDGFNEVSTF
SDMLGNDLGLSIAESNRARAGDETLFTRDAIANNKTMLSYGATASVEERIRNEGA
AGGSEGLERKLAIEYSSVESNOVNLDELVALRDAGVDIIDIILAEAGLAGEALYKAA
SDKIGFDALNKALKRATSEGGKRYGNATKAEAKTLAQAQOOTAKMSEKFLDIGKALE
PMNIGFEKVOQFIMVALIIVPLTFTVYTAFTVITLTKGLFVAVLAGKVOGKMLFKALT
TPFKLIDGIKIYVIGKLKELATYVSSGFIDRFKAPFPIIKWOKLIDMSKATYKIK
NLVFWKAEKEXKTSSEPERDKKFPDPNKTINNKMAEDYQRLQDEITNRQDITN
KTGARQEDALNLEKTIINENOKKEDYSKIFEDLTIDENKILIVGERSVNEFNNSV
DEVENYONLLEKESREBEITKTLPHDQVSALOKLDLEINERKAFVEXYKGFETL
NESNRQVVALLEKQVNEYKTAIDRSVEAQKALQKEITDLEMTMLPAKERASAK

KMAKIQAMYYKKFVDEHKSOFKELNETNNTIKOYAEKADOTTKSLYSDMIDGLNVEK
 NAFMKDIAGKFLNKDPTGSEIGEEFNHNLINGKVEKRNRIKAEKREDEDELEKREVELEK
 GAVRPMGEAVALELNGLDPPVWGLIKOGRKRIKAEKREDEDELEKREVELEK
 LEDRDEDELEKREKLESEIDDEYTEIEFLKQASQJSGEFOKRLHDVOTETKTK
 KDIETTOJLTKTEARKLEVERKNLSKLEPERIKQAEVDKNAEDMWLGPDRILK
 TKOILDEILKRIAKVSAGSIEIKLHAGFAVSNKPYMPNSWMLGPOPELIVR
 ITPAPIDENLRLEAKIIEAEITKLOKTSVNVNFYNNFNGVDLAEKLVMLKSK
 EHLMEFRMAE"
 complement(10776..11015)
 /gene="BBG11"
 complement(10776..11015)
 /note="BBG11"
 /note="hypothetical protein; identified by Glimmer;
 putative"

Query Match 13.1%; Score 98.4; DB 1; Length 29766;

Best Local Similarity 49.7%; Pred. No. 36-07; Matches 336; Conservative 0; Mismatches 331; Indels 9; Gaps 3;

QY 1 TGAATTTTCACCAATTAAGATGTGTTAAAGATGGAATGCTAATTTATGTTA 60
 1910 TGGCATACAAAAAATAATAGCATATTTCAATTTAAAGATTTATGAAAACTTT 1969
 61 AATCAAGAAGCTCTATTGGTAAGCAATTCAGATAACAATTTGAAAAAGTTAAATTA 120
 1970 ACATTAAATCTGCTTAGAAGCAAGAAATTCACAAATCTTTAGAAAAAGATTCA 2029
 QY 121 AATAGTCCAAAAACCTTTTAAATTCATTAATATGCTACATAATACAGTTAATA 180
 2030 ACTCCATTAATTAATCTAAATAATGTGCTAATTTGTTTATCAAACTTAATATTAG 2089
 QY 181 AAGG---GGTTTATGATTAAGTTTATTTATTTTGGTAATCTTTGCTTTT 237
 2090 GAGGTAAATTTATTAAGAAAAATTTTCATTAATTAATTTTATTTAGTAAATGC 2149
 QY 238 CTAGTTTCTCAGATGATTTCTAAAGCACTTTAATCTGGAGCGGAGAAAAATTT 297
 2150 AATCTTTCACAGATAAATCTGAAAAAGGTGTGAGT -GGAGATATCGCAGCTAA 2208
 QY 298 TGGTTATGAACATAAAGAAATCTCTGTACCATTTTATTTAGAACCTTTTATG 357
 2209 TGAATAATCAAGCAAAAAAGCAACCATCTACACCATTAATCTTTGAAATATCTTTTGT 2268
 QY 358 GGTTCGAGTAGGTTCTTTGCTCAAGGAGATATCTTGGAGGTCTTATCTTGAT 417
 2269 CTTTAGGATAGCATCTTTTGTTCAGAGATATTAATGTGTGCTGACATCTTGCT 2328
 QY 418 TTGATCGGCTGTATAGGGTTAATCTTACAGAGACTTATTTGACATCAAGATTG 477
 2329 CTCAGTGTGTTGAGAG-----AATACTTAATATGCTGACACATGATCTGAGAGTA 2383
 QY 478 ATTAATATCTTAAAAAAGCTGATTTTAAGTGACCTTGGGGTAAAGGAATGATGGCAG 537
 2384 TTGGAGCGCTTACAGAAACACAGCAAGTAAACCGAGGAGATATTATCAGAAATAG 2443
 QY 538 GTGTGTTCTATAGCTGTGACAAAGATGACAAATTTGTTCTCATTTACATTTGCTA 597
 2444 GAGGGCTCAATATGATGATCTTACATACATGGAATTAATTTCAATTTAAATTTGCTA 2503
 QY 598 AATAATTAACAGGAAGTGAATAATAGTCTTAATATAGCCTTGGAGATTTGAGCTTA 657
 2504 ATAGATACAGACGACCTTAAAAAAGGCTGGCATTTGCACTTGGCGGGTTGAACCCA 2563
 QY 658 GTTTTGAATTTAAT 673
 2564 ATTTGACATTTGAAT 2579
 DB

RESULT 6
 LOCUS AE001584 52971 bp DNA linear BCT 21-JUN-2002
 DEFINITION Borrelia burgdorferi plasmid lp56, complete plasmid sequence.
 ACCESSION AE001584

VERSION AE001584.1 GI:6382393
 KEYWORDS Lyme disease spirochete.
 SOURCE Borrelia burgdorferi
 ORGANISM Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia burgdorferi group.
 REFERENCE 1 (bases 1 to 52971)
 Casjens,S., Palmer,N., van Vugt,R., Huang,W.M., Stevenson,B., Rosa,P., Lathigra,R., Sutton,G., Peterson,J., Dodson,R.J., Haft,D., Hickey,E., Gwinn,M., White,O. and Fraser,C.M.
 A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi
 Mol. Microbiol. 35 (3), 490-516 (2000)
 JOURNAL 20138354
 MEDLINE 10672174
 PUBMED 10672174
 REFERENCE 2 (bases 1 to 52971)
 Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Lathigra,R., White,O., Dodson,R., Hickey,E.K., Gwinn,M., Peterson,J., van Vugt,R., Palmer,N., Haft,D., Rosa,P. and Stevenson,B.
 Direct Submission
 Submitted (11-NOV-1999) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
 location/Qualifiers
 1. 52971
 /organism="Borrelia burgdorferi"
 /db_xref="taxon:139"
 /plasmid="lp56"
 279..548
 /gene="BBQ01"
 279..548
 /gene="BBQ01"
 /note="similar to GB:U03641 PID:458212 percent identity: 67.53; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AA07704.1"
 /db_xref="GI:6382394"
 /translation="MTKYAIILVARKDAEICISGVNKNKOTENYKNLIVLLCLLF GYLGVRFYVGRIGGLYLTFEFLYGVLLIDLRIITNFKCN"
 710..802
 /gene="BBQ02"
 710..802
 /gene="BBQ02"
 /note="hypothetical protein; identified by Glimmer2; putative"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="AA07738.1"
 /db_xref="GI:6382428"
 /translation="MRKRIGICILVIEKSIIVGLMFKIVIR"
 856..1407
 /gene="BBQ03"
 856..1407
 /gene="BBQ03"
 /note="similar to GB:U31615 GB:U31418 PID:520783 PID:551742 percent identity: 97.81; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="outer membrane protein, putative"
 /protein_id="AA07705.1"
 /db_xref="GI:6382395"
 /translation="MRILGVFLIALALLGCLYLPDNOEAVOTEPENSESIDMSDE IYEGIFSSIKLYASEHRLVIEIKKLLISLKDPNRDVDPVADYNEFYENKRFIDGS ESKDLIKFGRYKNEQNKKFSEVYMLISCISELSPDKISGEGSPERYRPMRPR TAYQDLKVKREIERTINGKITL"
 1404..2268
 /gene="BBQ04"
 1404..2268
 /note="this region contains an authentic frame shift and is not the result of a sequencing artifact; similar to

gene

gene GP:320778 percent identity: 93.50; identified by sequence similarity: putative; protein p23, authentic frameshift
 2744. .3496
 /gene="BBQ05"
 2744. .3496
 /gene="BBQ05"
 /note="similar to GP:1553115 percent identity: 53.30; identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=1
 /product="antigen, p35, putative"
 /protein_id="AA07706.1"
 /db_xref="GI:6382396"
 /translation="MKYITCVFLLLNACNSDFSTQEDIKTPSDKEKSKSMNASSKEEDPNKIKTMTLNDLNLLEIANHEKEREKEMQEPDQGIILAFQBLDSVKI SEDPQOSKPKRYTSPSLAIDVAKMLDSEILNSIOIGLFINFRFGIGDSDIN HVYSKDLGLGLETLLDDPKNSPEKLTISTKEPFSKMLNLQLLDYKNDKDHRTENK LKSHRTAFEQDLKDEDAYPEPKNQDISLNL"
 3623. .4108
 /gene="BBQ06"
 3623. .4108
 /gene="BBQ06"
 /note="similar to GP:2239281 percent identity: 79.75; identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AA07707.1"
 /db_xref="GI:6382397"
 /translation="MQIRPAQKLEKGGVDATVVKYSKATITIAPIILNIFSLSIG SVAQGDYGGGALDSQVLGSLIMAGIMTGDICFVETSAIVTGGVLSIGGLTIA ASYIGIILPKRFANRYNADLKRRLGIALAGLEPNFDINGDSNYRLKNPIEQFTW G"
 complement(4249. .4986)
 /gene="BBQ07"
 complement(4249. .4986)
 /gene="BBQ07"
 /note="similar to PID:1143400 percent identity: 62.50; identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=1
 /product="conserved hypothetical protein"
 /protein_id="AA07708.1"
 /db_xref="GI:6382398"
 /translation="MGKKYKKRKEVLNRPFSHNNKELTINDNQDKLITNNYLKEK LKYNKDDIDNKEKMLIYELKEKELKYKDFSPDFQIFVIATQAVLYLKYVE KYLEASISDINKETGEPGRLGROLKENTLNKKNLESVNNKMSIRFPKKNKPEYF KCEFDKRACTILERLFTFKREYVLDLINYEYKRRKKNNKYLIFITIGAYILGMM MSFVLPQLSSOSKRFPGWGLSFLLLL"
 complement(5069. .5830)
 /gene="BBQ08"
 complement(5069. .5830)
 /gene="BBQ08"
 /note="similar to GB:D26185 SP:P37522 PID:467381 PID:860306 GB:AL009126 percent identity: 34.17; identified by sequence similarity: putative"
 /codon_start=1
 /transl_table=1
 /product="plasmid partition protein, putative"
 /protein_id="AA07709.1"
 /db_xref="GI:6382399"
 /translation="MDRKESNITITIASPKGVGKTYLTIFSYILKDLSKVLLIID PONSLSYFSKYIFNIDKCNSSYSLKDKDVGQCINKINDPISITPSHPILENNSEI LANYKOLLLEINRITNTVNFYDILDPNPGLTKNLSLWVTDYIILPOVEFSVEAE SLSISMTQINDIKDRNKSFNISITENQFLNRMDFKQVEELLYKVEYSIKGIIHYH NSVKVLINLELPSYKRYEETKNTLENTLNLF"
 complement(5803. .6339)
 /gene="BBQ09"
 complement(5803. .6339)
 /gene="BBQ09"
 /note="similar to GB:003641 PID:458217 percent identity: 58.67; identified by sequence similarity: putative"
 /codon_start=1

Db 3629 ATCTTGCACAGAGAACTTGAAAAAGTGTGGAGATATGCAACCTATGAATAT 3688
 Oy 306 GAACTAATAGAAAGATTCCTGTACATTTTATGACCTTTTATAGGTTCCGG 365
 Db 3689 GAAAGTAAAAAGCAACCATATACACCATATGACATCTTTTATCTTAGGA 3748
 Oy 366 ATAGTCTTTTGTCAAGGAGATATCCTTGAGGTCTCTATCTTGATTTGATGG 425
 Db 3749 ATGAGATCTTTGTCAAGGAGATATGAGTGTGGGACACTGTGGCTCAAGTG 3808
 Oy 426 GTTGTATAGGGTTAATACCTTACAGAGCTTATTTGACATCAAGATTTGATATAT 485
 Db 3809 CTGGAGGAATCTATATATGCTGTGATACAGCCGGGATTTGATCTTACAGAA 3868
 Oy 486 GCTAAAAAGCTGATTTTAACTGACCTTGGGTAAGGAGATGTTGGCAGGTGTGTT 545
 Db 3869 AGCAGAGCAACGATATTAACCGAGAGATATATCAGCAAT-----ACGAGGGCTC 3919
 Oy 546 ACTATGGCTGTGACAGATTTGACAGAAATGTTCTCCATTTACATTTGCTAATATAT 605
 Db 3920 ACAATGACGATCTCTACATCTGCAATATATCATTTCAATTAATTTGCTAATAGATAC 3979
 Oy 606 AACAGGAAGCTGAAAAATAGTCTTATATAGCCTTGGAGATTTGAGCCTAGTTTGTAT 665
 Db 3980 AACGAGACCTTAAAAAAGACTCGGCAATGCACTTGGCGGCTTGAACCAATTTTGCAC 4039
 Oy 666 ATTAACAT 673
 Db 4040 ATTGGAAT 4047

CDS
 /gene="BB101"
 174..608
 /gene="BB101"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB101"
 /protein_id="AAC6211.1"
 /db_xref="GI:2690108"
 /translation="MENQYEFNFIEFIKIIFYLYIMKSYKEQOEQIDHACRILI
 LTERFEILILENYSOKTLKKYENLNKKNLPSPNSTMKKYLKQLEKEIKIATK
 YEKNDOSLIYKNTLEKIMKLIELEPKELKQRIKNTT"
 736..1104
 /gene="BB102"
 736..1104
 /note="similar to GP.1655791 percent identity: 66.67;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAC66189.1"
 /db_xref="GI:2690086"
 /translation="MDGVINDPLVARMKKQIKENKLIILVKTLDHNNKELYSANK
 TNYVLIQNNRENAKTYQLRVNRTKILELIELEKRPYLRPTKNESEFGLY
 TLSPLEVCCTKIYMLHPNI"
 complement(1847..1972)
 /gene="BB103"
 complement(1847..1972)
 /gene="BB103"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB103"
 /protein_id="AAC66212.1"
 /db_xref="GI:2690109"
 /translation="WVYVLDNFSLYFGGLACVWDYYWVISLSEIVSFL"
 complement(2124..2219)
 /gene="BB104"
 complement(2124..2219)
 /gene="BB104"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB104"
 /protein_id="AAC66213.1"
 /db_xref="GI:2690110"
 /translation="WGRTIQRFLKRLKHLGLMRALPKIGESKYFIY"
 complement(2188..2310)
 /gene="BB105"
 complement(2188..2310)
 /gene="BB105"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB105"
 /protein_id="AAC66214.1"
 /db_xref="GI:2690111"
 /translation="MPYLVNWSNVQKETILKALKLKFQDRKSGNNPRTFN"
 2536..3351
 /gene="BB106"
 2536..3351
 /gene="BB106"
 /note="similar to GB.D26562 SP.P24247 GB.M83735 PID:473817
 PID:457107 percent identity: 33.62; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="pts protein (pts)"

RESULT 7
 AE000789/c
 LOCUS AE000789 27323 bp DNA linear BCT 17-DEC-1997
 DEFINITION Borrelia burgdorferi plasmid lp28-4, complete plasmid sequence.
 ACCESSION AE000789
 VERSION AE000789.1 GI:2690079
 KEYWORDS
 SOURCE
 ORGANISM
 Borrelia burgdorferi.
 Borrelia burgdorferi
 Bacteria; Spirochaetales; Spirochaetaceae; Borrelia; Borrelia
 burgdorferi group.
 1 (bases 1 to 27323)
 Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.,
 Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K.,
 Gwinn,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D., Richardson,D.,
 Peterson,J., Kerlavage,A.R., Quackenbush,J., Salzberg,S.,
 Hanson,M., van Vugt,R., Palmer,N., Adams,M.D., Gocayne,J.,
 Venter,J.C. et al.
 Genomic sequence of a Lyme disease spirochaete, Borrelia
 burgdorferi
 Nature 390 (6660), 580-586 (1997)
 JOURNAL MEDLINE 98065943
 PUBMED 9403685
 REFERENCE
 2 (bases 1 to 27323)
 Fraser,C.M., Casjens,S., Huang,W.M., Sutton,G.G., Clayton,R.A.,
 Lathigra,R., White,O., Ketchum,K.A., Dodson,R., Hickey,E.K.,
 Gwinn,M., Dougherty,B., Tomb,J.F., Fleischmann,R.D.,
 Richardson,D., Peterson,J., Kerlavage,A.R., Quackenbush,J.,
 Salzberg,S., Hanson,M., van Vugt,R., Palmer,N., Adams,M.D.,
 Gocayne,J.D., Weidman,J., Utterback,T., Matthey,L., McDonald,L.,
 Artlieb,P., Bowman,C., Garland,S., Fujii,C., Cotton,M.D., Horst,K.,
 Roberts,K., Hatch,B., Smith,H.O. and Venter,J.C.
 Direct Submission
 Submitted (08-DEC-1997) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 TITLE
 JOURNAL
 FEATURES
 SOURCE
 1..27323
 /organism="Borrelia burgdorferi"
 /db_xref="taxon:139"
 /plasmid="lp28-4"
 174..608
 gene

gene
 CDS
 /gene="BB101"
 174..608
 /gene="BB101"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB101"
 /protein_id="AAC6211.1"
 /db_xref="GI:2690108"
 /translation="MENQYEFNFIEFIKIIFYLYIMKSYKEQOEQIDHACRILI
 LTERFEILILENYSOKTLKKYENLNKKNLPSPNSTMKKYLKQLEKEIKIATK
 YEKNDOSLIYKNTLEKIMKLIELEPKELKQRIKNTT"
 736..1104
 /gene="BB102"
 736..1104
 /note="similar to GP.1655791 percent identity: 66.67;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAC66189.1"
 /db_xref="GI:2690086"
 /translation="MDGVINDPLVARMKKQIKENKLIILVKTLDHNNKELYSANK
 TNYVLIQNNRENAKTYQLRVNRTKILELIELEKRPYLRPTKNESEFGLY
 TLSPLEVCCTKIYMLHPNI"
 complement(1847..1972)
 /gene="BB103"
 complement(1847..1972)
 /gene="BB103"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB103"
 /protein_id="AAC66212.1"
 /db_xref="GI:2690109"
 /translation="WVYVLDNFSLYFGGLACVWDYYWVISLSEIVSFL"
 complement(2124..2219)
 /gene="BB104"
 complement(2124..2219)
 /gene="BB104"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB104"
 /protein_id="AAC66213.1"
 /db_xref="GI:2690110"
 /translation="WGRTIQRFLKRLKHLGLMRALPKIGESKYFIY"
 complement(2188..2310)
 /gene="BB105"
 complement(2188..2310)
 /gene="BB105"
 /note="hypothetical protein; identified by Glimmer;
 putative"
 /codon_start=1
 /transl_table=11
 /product="B. burgdorferi predicted coding region BB105"
 /protein_id="AAC66214.1"
 /db_xref="GI:2690111"
 /translation="MPYLVNWSNVQKETILKALKLKFQDRKSGNNPRTFN"
 2536..3351
 /gene="BB106"
 2536..3351
 /gene="BB106"
 /note="similar to GB.D26562 SP.P24247 GB.M83735 PID:473817
 PID:457107 percent identity: 33.62; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=11
 /product="pts protein (pts)"

gene
CDS
/protein_id="AAC66190.1"
/db_xref="GI:2690087"
/translation="MKIKNTANILIFELSYLVNMGWESKINFPNENSQSNILISAT
KAEIEIKILLQNNKYSIEEHRKKKRTIAGKLMNDHITITATGVKINTALPMSTII
SKYISHLINGNAGSVISDRNKPILIGDVAISTETYSFDLDRFGEYIEGHPENR
KEFNATALIRTSKIKINNTITSYMGILITGDOFIDHOTFEIPEEPENALAIWESMA
MAOVAYGKIPFIIRIGSDIVNENNVDDYKFLKASSSAKIVENLIKLM"
complement(3343. .3441)
/gene="BB107"
complement(3343. .3441)
/gene="BB107"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BB107"
/protein_id="AAC66215.1"
/db_xref="GI:2690112"
/translation="WQLKGAIIKLEDYEKIKRIQLNLIKLIH"
3576. .3755
/gene="BB108"
3576. .3755
/gene="BB108"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BB108"
/protein_id="AAC66196.1"
/db_xref="GI:2690093"
/translation="MISFPDNPSSASNIFKITTPCKLAKKLIPLIQRKQFNNESS
FYKNKIRYSLNIGK"
3745. .3882
/gene="BB109"
3745. .3882
/gene="BB109"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BB109"
/protein_id="AAC66197.1"
/db_xref="GI:2690094"
/translation="MASEAKIDAESKNKEVNSQKKSNKSNIEVKDTLGLVKAIKKS
I"
3911. .4315
/gene="BB110"
3911. .4315
/gene="BB110"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BB110"
/protein_id="AAC66198.1"
/db_xref="GI:2690095"
/translation="MGRATTPHOQKLVNLENGLKMINLDELKTAIDGADNNISKKK
NLRKYVDNENQNSLIKIVYLNKSSNMEKIEKKKIKITLMHVEYIEFSDVDELKKKK
KNEHSILITLNEAIGKIKNSAIAIHVCENID"
complement(4623. .4721)
/gene="BB111"
complement(4623. .4721)
/gene="BB111"
/note="hypothetical protein; identified by Glimmer;
putative"
/codon_start=1
/transl_table=1
/product="B. burgdorferi predicted coding region BB111"
/protein_id="AAC66199.1"
/db_xref="GI:2690096"
/translation="MSKLOGYSTVNVVFPKIKIKILIVNKGSVLF"
5128. .5346
/gene="BB112"

[illegible]

ACCESSION	AA4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
VERSION	AC117070.1
KEYWORDS	HTG; HTGS_PHASE2.
SOURCE	Dictyostelium discoideum.
ORGANISM	Dictyostelium discoideum
REFERENCE	Eukaryota: Mycetozoa: Dictyostelidae: Dictyostelium.
AUTHORS	1 (bases 1 to 156533) Gloeckner, G., Eichinger, L., Szafranski, K., Pachebat, J., Dear, P., Lehmann, R., Baumgartner, C., Parra, G., April, J. F., Guigo, R., Kumpf, K., Tüngel, A. B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A. and Noegel, A. A.
TITLE	Sequence and Analysis of Chromosome 2 of Dictyostelium
JOURNAL	Unpublished
REMARK	The Dictyostelium Genome Sequencing Consortium
AUTHORS	2 (bases 1 to 156533) Baumgart, C.
JOURNAL	Direct Submission
COMMENT	Submitted (06-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany CDS predictions from GeneID may contain errors. Further Information is available from IMB Jena, Department of Genome Analysis (http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I (http://www.uni-koeln.de/dictyostelium/project.shtml) Funding Agency : Deutsche Forschungsgemeinschaft (DFG). * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved.
FEATURES	Location/Qualifiers 1. 156533 /organism="Dictyostelium discoideum" /strain="Ax4" /db_xref="taxon:44689" /chromosome="2" /map="2268371-2255204" join(3827. 3945,5512. 5557) /note="ORF ID:dd_00443" /codon_start=1 /product="hypothetical protein" /protein_id="AA09328.1" /db_xref="GI:20066214" /translation="MSADQPLRLTMLMSADQPLPWQISQIMSVPTLLITHSYGILASIDEGNOTA" complement(7554. 8315) /note="ORF ID:dd_00439" /codon_start=1 /product="DiscoIdin I, D chain" /protein_id="AA09325.1" /db_xref="GI:20066211" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" complement(9039. 9800) /note="ORF ID:dd_00438" /codon_start=1 /product="DiscoIdin I, C chain and B chain." /protein_id="AA09326.1" /db_xref="GI:20066212" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" join(12482. 12692,13544. 13672,14278. 14315) /note="ORF ID:dd_00437" /codon_start=1 /product="hypothetical protein" /protein_id="AA09327.1"
CDS	
FEATURES	Location/Qualifiers 1. 156533 /organism="Dictyostelium discoideum" /strain="Ax4" /db_xref="taxon:44689" /chromosome="2" /map="2268371-2255204" join(3827. 3945,5512. 5557) /note="ORF ID:dd_00443" /codon_start=1 /product="hypothetical protein" /protein_id="AA09328.1" /db_xref="GI:20066214" /translation="MSADQPLRLTMLMSADQPLPWQISQIMSVPTLLITHSYGILASIDEGNOTA" complement(7554. 8315) /note="ORF ID:dd_00439" /codon_start=1 /product="DiscoIdin I, D chain" /protein_id="AA09325.1" /db_xref="GI:20066211" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" complement(9039. 9800) /note="ORF ID:dd_00438" /codon_start=1 /product="DiscoIdin I, C chain and B chain." /protein_id="AA09326.1" /db_xref="GI:20066212" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" join(12482. 12692,13544. 13672,14278. 14315) /note="ORF ID:dd_00437" /codon_start=1 /product="hypothetical protein" /protein_id="AA09327.1"
CDS	
FEATURES	Location/Qualifiers 1. 156533 /organism="Dictyostelium discoideum" /strain="Ax4" /db_xref="taxon:44689" /chromosome="2" /map="2268371-2255204" join(3827. 3945,5512. 5557) /note="ORF ID:dd_00443" /codon_start=1 /product="hypothetical protein" /protein_id="AA09328.1" /db_xref="GI:20066214" /translation="MSADQPLRLTMLMSADQPLPWQISQIMSVPTLLITHSYGILASIDEGNOTA" complement(7554. 8315) /note="ORF ID:dd_00439" /codon_start=1 /product="DiscoIdin I, D chain" /protein_id="AA09325.1" /db_xref="GI:20066211" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" complement(9039. 9800) /note="ORF ID:dd_00438" /codon_start=1 /product="DiscoIdin I, C chain and B chain." /protein_id="AA09326.1" /db_xref="GI:20066212" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" join(12482. 12692,13544. 13672,14278. 14315) /note="ORF ID:dd_00437" /codon_start=1 /product="hypothetical protein" /protein_id="AA09327.1"
CDS	
FEATURES	Location/Qualifiers 1. 156533 /organism="Dictyostelium discoideum" /strain="Ax4" /db_xref="taxon:44689" /chromosome="2" /map="2268371-2255204" join(3827. 3945,5512. 5557) /note="ORF ID:dd_00443" /codon_start=1 /product="hypothetical protein" /protein_id="AA09328.1" /db_xref="GI:20066214" /translation="MSADQPLRLTMLMSADQPLPWQISQIMSVPTLLITHSYGILASIDEGNOTA" complement(7554. 8315) /note="ORF ID:dd_00439" /codon_start=1 /product="DiscoIdin I, D chain" /protein_id="AA09325.1" /db_xref="GI:20066211" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" complement(9039. 9800) /note="ORF ID:dd_00438" /codon_start=1 /product="DiscoIdin I, C chain and B chain." /protein_id="AA09326.1" /db_xref="GI:20066212" /translation="MSTGSLVTLILGNACHLRTSTNYNDVHTQFNAVLNKNKNTNIDSEACSSIVDTNQTIVAGCEVPRTFPCVALOGGDHDQWTSTKIRYSLDNTWSEYKGAALITGVDRNTVNHFDPTPIRARSIALHPLTNNHLSLRCEFTQVQSSVQVQADITGNCALNTGSGKREVVVPKQFQEPATLPKVALNFQDIDCDTADNTQIRIGVOPANITTKGDCVFPTMANKYVSLRADYIATALE" join(12482. 12692,13544. 13672,14278. 14315) /note="ORF ID:dd_00437" /codon_start=1 /product="hypothetical protein" /protein_id="AA09327.1"
CDS	
FEATURES	Location/Qualifiers 1. 156533 /organism="Dictyostelium discoideum" /strain="Ax4" /db_xref="taxon:44689" /chromosome="2" /map="2268371-2255204" join(3827. 3945,5512. 5557) /note="ORF ID:dd_00443" /codon_start=1 /product="hypothetical protein" /protein_id="AA09328.1" /db_xref="GI:20066214" /translation="MSADQPLRLTMLMSADQPLPWQISQIMSVPTLLITHSYGILASIDEGNOTA" complement(7554. 8315) /note="ORF ID:dd_

LOCUS	AX251990	AX251990	Sequence 251 from Patent WO0168911.	AX251990	AX251990	AX251990.1	GI:15985345
DEFINITION	Sequence 251 from Patent WO0168911.						
ACCESSION	AX251990						
VERSION	AX251990.1						
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
BASE COUNT	1749	a	289	c	1822	g	3250
ORIGIN							
Query Match			8.8%; Score 66; DB 6; Length 7110;				
Best Local Similarity			51.0%; Pred. No. 0.08;				
Matches	156;	Conservative	0;	Mismatches	150;	Indels	0;
							Gaps 0;
OY	35	AATTGAAATGCTAATTTTATGTTTAAATCAAGAAGCTCTATTGGTTAAGCGAATTCGAG	94				
Db	1672	AATTGAAATTTTAAAAAGATATTATTTTTGAATTTTATTAGTTAGATAGGTTAGTTA	1731				
OY	95	TACAAATTTGAAAAAGTTAATTTAATTAATAGTCCAAAAGCCTTTTAAATTCATTAA	154				
Db	1732	TTTGGGTGAAAAAGAAAAAATTAATAAATTGAAAAAGATGATTTAGTATTTT	1791				
OY	155	TATGCTACCAATTAATCCAGTTTAATTAATGAAGGGTTTATGAATAGTTTAAATTTAT	214				
Db	1792	AGAAATTAATTAATTAATTAATGAATTAACAGAGAGATTTTATTTAATTTTAAATTAAT	1851				
OY	215	TTTGGTAAATCTTTTGTGCTTTTCTAGCTTTTGCTCAAGATGATTCGTAAAGCATTTTAA	274				
Db	1852	TTGATGTGTGTTGGTGGTTGAATTAATTTGTTAATTTGATGATTTTATTTTAAATTA	1911				
OY	275	TCTGGAGCGGAGAAAAATTTTGGTTTATGAACCTAATAAGAAAGATCTCTGTAC	334				
Db	1912	TTAGAGACAGATGAGGATTTGATGTTAGAGAAAGTGAATAATTTTATTTTATTTT	1971				
OY	335	ATTTT 340					
	1972	TTTTT 1977					
RESULT 11							
LOCUS	AX346458	AX346458	7110 bp	DNA	linear	PAT 01-FEB-2002	
DEFINITION	Sequence 1529 from Patent WO0200928.						
ACCESSION	AX346458						
VERSION	AX346458.1						
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							
Location/Qualifiers							
1..7110							
/organism="synthetic construct"							
/db_xref="taxon:32630"							
/note="Chemically treated genomic DNA (Homo sapiens)"							

BASE COUNT	1749 a	289 c	1822 g	3250 t	ORIGIN
Query Match	8.8%;	Score 66;	DB 6;	Length 7110;	
Best Local Similarity	51.0%;	Pred. No. 0.08;			
Matches 156;	Conservative 0;	Mismatches 150;	Indels 0;	Gaps 0;	
QY	35	AATTGAATTCGTAATTTATGCTTAATTCAGAAAGCTCTATTGGTAAGCGAATTTCCAG	94		
Db	1672	AATTGAATTTTAAAAAGATATTTTTCGAAATTTTATTTAGTTAGATAGGCTAGTTA	1731		
QY	95	TAAACAATTTGAAAAAGTTAAATTTAAATTAATAGTCCAAAAAGCCTTTTAAATTCATTAA	154		
Db	1732	TTTGGGTGAAAGGAAAGAAAGAAATTAATAATTTGAAAGGATGTAATGTAATTTT	1791		
QY	155	TATGCTACCATTAATACCAAGTTAATTAAGAGGGCTTTTATGATTAAGATTTTAATTTTAT	214		
Db	1792	ACAATTAATTTAAATTTAGCAATAGACAGAGATTTTATTTAATTTTAATTAATTAAT	1851		
QY	215	TTTGGTAATCTTTTGCTTTTCTTAGATTTTGCTCAAGATGATTCATAAGCATTTTAA	274		
Db	1852	TTGATGTTGTTTGGGTGGTTAAATATTTGTAATATGATGATTTTTTATTTATTTA	1911		
QY	275	TCTGGAGCGGAGAGAAATTTTGGTTATGAACCTATATGAAGATTCCTGTGACC	334		
Db	1912	TTAGAGAGAGATGAGAGGATTTGATGTTTAGAAGATGTAATAATTTATTTTATTTT	1971		
QY	335	ATTTT 340			
Db	1972	TTTTT 1977			
RESULT 12					
LOCUS	AX349019	7110 bp	DNA	Linear	PAT 06-FEB-2002
DEFINITION	Sequence 477 from Patent WO0202807.				
ACCESSION	AX349019				
VERSION	AX349019.1	GI:18615054			
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE	1	synthetic construct.			
AUTHORS	Olek,A., Piepenbrock,C. and Berlin,K.				
TITLE	Diagnosis of diseases associated with cell signalling				
JOURNAL	Patent: WO 0202807-A 477 10-JAN-2002;				
FEATURES	Epigenomics AG (DE)				
source	Location/Qualifiers				
	1..7110				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note="chemically treated genomic DNA (Homo sapiens)"				
BASE COUNT	1749 a	289 c	1822 g	3250 t	
ORIGIN					
Query Match	8.8%;	Score 66;	DB 6;	Length 7110;	
Best Local Similarity	51.0%;	Pred. No. 0.08;			
Matches 156;	Conservative 0;	Mismatches 150;	Indels 0;	Gaps 0;	
QY	35	AATTGAATTCGTAATTTATGCTTAATTCAGAAAGCTCTATTGGTAAGCGAATTTCCAG	94		
Db	1672	AATTGAATTTTAAAAAGATATTTTTCGAAATTTTATTTAGTTAGATAGGCTAGTTA	1731		
QY	95	TAAACAATTTGAAAAAGTTAAATTTAAATTAATAGTCCAAAAAGCCTTTTAAATTCATTAA	154		
Db	1732	TTTGGGTGAAAGGAAAGAAAGAAATTAATAATTTGAAAGGATGTAATGTAATTTT	1791		
QY	155	TATGCTACCATTAATACCAAGTTAATTAAGAGGGCTTTTATGATTAAGATTTTAATTTTAT	214		
Db	1792	ACAATTAATTTAAATTTAGCAATAGACAGAGATTTTATTTAATTTTAATTAATTAAT	1851		
QY	215	TTTGGTAATCTTTTGCTTTTCTTAGATTTTGCTCAAGATGATTCATAAGCATTTTAA	274		
Db	1852	TTGATGTTGTTTGGGTGGTTAAATATTTGTAATATGATGATTTTTTATTTATTTA	1911		
QY	275	TCTGGAGCGGAGAGAAATTTTGGTTATGAACCTATATGAAGATTCCTGTGACC	334		
Db	1912	TTAGAGAGAGATGAGAGGATTTGATGTTTAGAAGATGTAATAATTTATTTTATTTT	1971		
QY	335	ATTTT 340			
Db	1972	TTTTT 1977			


```

BASE COUNT      2133 a      68 c      1273 g      4005 t
ORIGIN
/note="chemically treated genomic DNA (Homo sapiens)"

```

Query Match	8.5%	Score	63.8	DB	6	Length	7479
Best Local Similarity	46.9%	Pred. No.	0.18				
Matches	234	Conservative	0	Mismatches	262	Indels	3
						Gaps	1

OY	17	TAAAGATATGTGTAAGAAAGAAATGAAATTCGTAATTTATGGTTAAATCAAGAGCTCAT	76
Db	6369	TGAATATTTTTTTTTTACGTATTTTCTTTTACGTTTGTTTTAGCTTAGCTTCGTTT	6428
OY	77	TGCTAACCGCAATTTCCAGTACACATTTGAAAAAGTTAAATTTAAATGCTCCAAAGCC	136
Db	6429	TAGGTAAAGATATTTTTTTTTTAAATTTATATATTTTTTATATTTTATATTTTATAGTA	6488
OY	137	TTTTTTTAAATTCATTA--ATATGCTACCATTAATCCAGTTTAAATAAGGGGTGTTAT	193
Db	6489	TTTAAATATATATTTATAGATTTTTTTTAAAGCTTTTAATAGCTTTTAAATTTTTTTTAT	6548
OY	194	GAAATACGTTTTTAATTTTATTTTGGTAATCTTTTGCTGCTTTTCTAGCTTTGCTCAAGA	253
D	6549	TTTTTCGATTTTTTATTTTGAATATGTAAATTTTTTTTTTTTTCGTATTTGAGTAAATTT	6608
OY	254	TGATTTCAAAGACCTTTTATATCTGGAGCGGAGAAAAATTTTGGTTTTATGAACTAA	313
Db	6609	TATTCGTTGGCAATTTTTTATTTATTTTCGGTTATAGAGTATTATATTTTATTTATATAT	6668
OY	314	TAAAGAAATTCCTCTGTACCATTTTTTATTTGAACCTTTTTTTTAAAGGTTCCGGATAGCTC	373
Db	6669	TTTTTAAAGACTTATTTTCGTTTTTATTTATTAAGAAGTATGATTTTATTTTAAAGTGT	6728
OY	374	TTTTGTCACAGAGATATCTCTGGAGGTCCTATATCTTGATTTGATGGCGTGGCAT	433
Db	6729	TATAGCTTTAAAGTTTATTTATTTTTTTTTTTTTTTTATTAAGTATATTTATTTTGTTAATAA	6788
OY	434	AGGCTTATATCTTACAGAGACTTTTGGACATCAAGATTTTGCATTAATATGCTAAAAA	493
Db	6789	TTTTTTTTTAAATATATATAGCTTTATTTGTAAAAAGTTAGATGATTTATTTATTTAGT	6848
OY	494	AGCTGATTTTAAGTGAGCT	512
Db	6849	TTAGAGATTTTAAAGTTATTT	6867

Search completed: March 27, 2003, 03:12:31
Job time : 1747.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 00:05:42 ; Search time 165.602 Seconds
(without alignments)
10185.565 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749
Sequence: 1 tagaatttcacaataaa.....attatctagaataatgggtg 749

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	749	100.0	749	20	AAX30100
2	607	81.0	862	20	AAX30099
3	558.2	74.5	759	20	AAX30098
4	402.8	53.8	540	20	AAX61505
5	368.8	49.2	477	20	AAX61506
6	98.4	13.1	1761	20	AAX20310
7	91.6	12.2	9399	20	AAX20261
8	82.2	11.0	7074	20	AAX20266
9	66	8.8	7110	22	AAS46489

10	66	8.8	7110	24	ABL70587	Chemically treated
11	66	8.8	7110	24	ABL33556	Human immune syste
12	66	8.8	7110	24	AAS61290	Human immune syste
13	66	8.8	34688	24	ABQ67059	Human angiotensin
14	63.8	8.5	7441	24	ABK40058	Human chemically p
15	63.8	8.5	7479	24	AAS63345	Human chemically p
16	63.8	8.5	14095	24	ABL32476	Human immune syste
17	63.2	8.4	6131	24	ABL32891	Human immune syste
18	62.2	8.3	6391	24	ABK28137	DNA transcription
19	61.8	8.3	14023	24	ABL34105	Human immune syste
20	60.2	8.0	9707	24	ABL33421	Human immune syste
21	59	7.9	3873	24	ABQ67089	Human angiotensin
22	59	7.9	5675	24	ABL33276	Human immune syste
23	59	7.9	6464	24	ABL33515	Human immune syste
24	58.6	7.8	9731	24	ABL32890	Human immune syste
25	58.4	7.8	8237	22	AAS46802	Tumour suppressor
26	58	7.7	15387	24	ABL32184	Human immune syste
27	57.6	7.7	7781	24	ABL33169	Human immune syste
28	57.2	7.6	13125	24	ABL70283	Chemically treated
29	57.2	7.6	13125	24	ABL33226	Human immune syste
30	57.2	7.6	13125	24	ABL34556	Human immune syste
31	57	7.6	4163	24	ABL34330	Human metastasis a
32	57	7.6	7348	22	AAS46336	Human immune syste
33	56.8	7.6	6244	24	ABL32484	Tumour suppressor
34	56.8	7.6	8077	24	ABL33176	Human immune syste
35	56.6	7.6	9021	24	ABK33988	Human immune syste
36	56.6	7.6	9021	24	ABK40075	Human DNA for stag
37	56.6	7.6	9021	24	ABL34232	Human chemically p
38	56.4	7.5	19380	24	AAS61426	Human immune syste
39	56	7.5	6022	22	AAS46661	Human gene regulat
40	56	7.5	17594	24	ABL34026	Tumour suppressor
41	55.4	7.4	14147	22	AAS46743	Human immune syste
42	55.4	7.4	14147	24	ABK33955	Tumour suppressor
43	55.2	7.4	6169	22	AAS46370	Human DNA for stag
44	55.2	7.4	6169	24	ABN80097	Tumour suppressor
45	55.2	7.4	17538	24	ABL33156	Human chemically m

ALIGNMENTS

RESULT 1	
AAX30100	standard; DNA; 749 BP.
ID	AAX30100;
AC	AAX30100;
XX	
XX	
DT	17-JUN-1999 (first entry)
DE	Borrelia garinii IP90 protein encoding cDNA.
XX	
KW	Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.
XX	
OS	Borrelia garinii.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	192..725
FT	/*tag= a
XX	
PN	W09912960-A2.
XX	
PD	18-MAR-1999.
XX	
PF	04-SEP-1998; 98WO-IB01424.
XX	
PR	16-SEP-1997; 97US-0059036.
PR	10-SEP-1997; 97DK-0001041.
XX	
PA	(SYMB-) SYMBICOM AB.
XX	
PI	Bergstroem S;
XX	
DR	WPI; 1999-215027/18.

DR P-PSDB: AAY04280.
 XX Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
 PT protein P13
 XX
 PS Claim 1; Page 113-114; 118pp; English.
 XX
 CC The present sequence encodes a *Borrelia garinii* IP90 protein. The
 CC present invention describes an isolated nucleic acid (I) that: (i)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within *B. burgdorferi*
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.
 CC
 XX
 XX
 SQ Sequence 749 BP; 234 A; 82 C; 150 G; 283 T; 0 other;
 Query Match 100.0%; Score 749; DB 20; Length 749;
 Best Local Similarity 100.0%; Pred. No. 2; 2e-134;
 Matches 749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 601 ATTATACAGAGAGCTGAATAAAGTCTTAATATAGCCTTGAGAGATTGACCTACTT 660
 QY 661 TTGATATTATACATGGGCCACAGCTAGTCTTGCTTTGGACTGCTTTCAAAAAAGCT 720
 DB 661 TTGATATTATACATGGGCCACAGCTAGTCTTGCTTTGGACTGCTTTCAAAAAAGCT 720
 QY 721 ATTAATTTTATTTATCTAGAAAAATGGGTG 749
 DB 721 ATTAATTTTATTTATCTAGAAAAATGGGTG 749
 RESULT 2
 AAX30099
 ID AAX30099 standard; DNA; 862 BP.
 AC AAX30099;
 DT 17-JUN-1999 (first entry)
 XX
 XX *Borrelia afzelii* ACAC protein encoding cDNA.
 DE
 XX *Borrelia*; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.
 XX
 XX *Borrelia afzelii*.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 219..755
 FT /*tag= a
 FT
 XX W09912960-A2.
 XX
 XX 18-MAR-1999.
 XX
 XX 04-SEP-1998; 98W0-1B01424.
 XX
 XX 16-SEP-1997; 97US-0059036.
 ER 10-SEP-1997; 97DK-0001041.
 XX
 XX (SYMB-) SYMBICOM AB.
 PA
 XX
 PI Bergstroem S;
 XX
 DR WPI; 1999-215027/18.
 DR P-PSDB: AAY04279.
 PT Nucleic acid from *Borrelia burgdorferi* encoding virulence associated
 PT protein P13
 PS
 PS Claim 1; Page 110-111; 118pp; English.
 XX
 CC The present sequence encodes a *Borrelia afzelii* ACAC protein. The
 CC present invention describes an isolated nucleic acid (I) that: (i)
 CC encodes a polypeptide fragment (II) immunologically reactive with
 CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
 CC *Borrelia burgdorferi* B313, but not reactive with most proteins from
 CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
 CC to specific nucleic acid sequences but not to genomic DNA from most
 CC other spirochaetes. (I), (II) and transformed cells are useful in
 CC vaccines to protect against infection by *B. burgdorferi* sensu lato.
 CC (I), (II) and antibodies raised against (II) are used to diagnose such
 CC infections, in standard immunoassays or amplification/hybridization
 CC tests. (I) are also used to produce recombinant (II). The 13 kD protein
 CC is involved in virulence and is highly conserved within *B. burgdorferi*
 CC sensu lato, but is absent from other *Borrelia* species (e.g. those
 CC responsible for relapsing fever or avian borreliosis). It should provide
 CC a more specific and sensitive antibody response, and diagnosis,
 CC compared with use of whole bacteria as antigen.
 CC
 XX
 XX
 SQ Sequence 862 BP; 265 A; 94 C; 162 G; 341 T; 0 other;
 Query Match 81.0%; Score 607; DB 20; Length 862;
 Best Local Similarity 90.1%; Pred. No. 3e-107;
 Matches 673; Conservative 0; Mismatches 70; Indels 4; Gaps 2;

```
OY 1 TAGAATTTTCAACAATTAAGATATTGTTAAAGAAATGAAATGCTAATTTATGCTA 60
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 27 TAGAATTTTCCGACAAATAAAGACATATTATTAAGAAATGAAATGCTAATTTATGCTA 86
OY 61 AATCAAGAAGCTCTATGTTGAAGCAATTCGAGTAACAAATTTGAAAAAGTTAAATTTA 120
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 87 AATCAAGAAGCTCTATGTTGAAGCAATTTCAAGTAATGCTTTGAAAAAGTTAAATTTA 146
OY 121 AATAGTCCAAAAGCTTTTAAATTTTCAATTAATGCTACCAATATACGAG-TTTAAT 179
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 147 AATAGTCTTAAACACCTTTTAAATTTTCAATTAATGCTACCAATATACGAGTTTAAAT 206
OY 180 AAGGGGTTTATGTAATGTAATTTTATTTATGCTATGCTATGCTATGCTATGCTATGCT 239
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 207 AATAGGTTTATGTAATTAATTTTATTTATGTTGCTATGCTATGCTATGCTATGCTATGCT 266
OY 240 AGTTTTCCTCAA--GAGATTTCTAAAGCACTTTTAACTCTGGAGCGGAGAAAAATTT 296
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 267 AGCTTTGCTCAAGCTGATGATTTCTAAAGCGCTTTTAAATTTGGAGCGGAGAAAAACTT 326
OY 297 TTGGTTTATGAACATTAATAGAAAGATTCCTTTGACATTTTATTTATGAACTTTTATTA 356
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 327 TTAGCTTATGAACATTAATAGAAAGATTCCTATTTGCTCCATTTTATTTGAACCTTTTATTA 386
OY 357 GGGTTCGGGATAGTTCTTTTCTGCTCAAGAGATATCCCTTGGAGGTTCTCTTATTTCTTGA 416
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 387 GGGTTTGGAAATAGTTCTTTTCTGCTCAAGAGATATTTCTTGGGGTTCCTTATTTCTTGA 446
OY 417 TTGATGCGGTTGATATAGGGTTAATACTTACAGAGCTTAAATTTGACATCAAGATTTT 476
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 447 TTGATGCGGTTGATATAGGGTTAATACTTACAGAGCTTAAATTTGATCAAGCTCTT 506
OY 477 GATTAATATGCTAAAAAGCTGATTTTAAAGTGAAGCTTTGGGTTAAGGAAATGATGTTGCA 536
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 507 GATTAAGATGCTCAAAAGCCCTTTTAAAGTGAAGCTTTGGGTTAAGGAAATGATGTTGCA 566
OY 537 GGTGGTTCATATGCGCTGACAGATGACAGAAATGCTCTCCATTTACATTTGCT 596
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 567 GGTGAGTTTACTATGCTGCTGACAGATGACAGAAATTTATTTATTCCTTATTTGCT 626
OY 597 AATTAATTAACAGAGCTGAAAAATAGCTTAATATAGCTTGGAGAGATTTGAGCCT 656
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 627 AATAGTATTAATAGGAACCTGAAAAATAGCTTAATATAGCTTTTGGAGGTTTGGCCT 686
OY 657 AGTTTGAATTAACATGCGCAAGCTAGCTGCTGTTTGGAGCTGCTTTCAAAAAA 716
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 687 AGTTTGAATTAATATGCGCAAGCTAGCTGCTGTTTGGAGCTGCTTTCAAAAAA 746
OY 717 AGCTAATTAATTAATTAATCTAGAAAA 743
   ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 747 AGCTAATTAATTAATTAATTAATTA 773
```

RESULT 3
AAK30098
ID AAK30098 standard; DNA; 759 BP.

AAK30098;
17-JUN-1999 (first entry) .

Borrelia burgdorferi B31 protein encoding cDNA.

Borrelia; P13 antigen; spirochaete; vaccine; infection; diagnosis; ss.

Borrelia burgdorferi.

Key Location/Qualifiers

FT CDS 170..709 /tag- a

PN W09912960-A2.

```
XX 18-MAR-1999.
PD 04-SEP-1998; 98WO-1B01424.
XX
PE 16-SEP-1997; 97US-0059036.
XX 10-SEP-1997; 97DK-0001041.
XX
PA (SYMB-) SYMBICOM AB.
XX
PI Bergstroem S;
XX WPI: 1999-215027/18.
XX P-PSDB: AAV04278.
DR
XX
XX Nucleic acid from Borrelia burgdorferi encoding virulence associated
PT protein P13
XX
XX Claim 1; Page 107-108; 118pp; English.
XX
CC The present sequence encodes a Borrelia burgdorferi B313 protein. The
CC present invention describes an isolated nucleic acid (I) that: (i)
CC encodes a polypeptide fragment (II) immunologically reactive with
CC rabbit polyclonal antibody raised against a 13 kDa polypeptide of
CC Borrelia burgdorferi B313, but not reactive with most proteins from
CC other spirochaetes; and/or (ii) hybridizes under stringent conditions
CC to specific nucleic acid sequences but not to genomic DNA from most
CC other spirochaetes. (I), (II) and transformed cells are useful in
CC vaccines to protect against infection by B. burgdorferi sensu lato.
CC (I), (II) and antibodies raised against (II) are used to diagnose such
CC infections. In standard immunoassays or amplification/hybridization
CC tests, (I) are also used to produce recombinant (II). The 13 kD protein
CC is involved in virulence and is highly conserved within B. burgdorferi
CC sensu lato, but is absent from other Borrelia species (e.g. those
CC responsible for relapsing fever or avian borreliosis). It should provide
CC a more specific and sensitive antibody response, and diagnosis,
CC compared with use of whole bacteria as antigen.
SQ Sequence 759 BP; 228 A; 90 C; 157 G; 284 T; 0 other;
Query Match 74.5%; Score 558.2; DB 20; Length 759;
Best Local Similarity 87.0%; Pred. No. 6; 3e-98;
Matches 638; Conservative 0; Mismatches 88; Indels 7; Gaps 2;
```

Db 421 GCTTATCTGCGGGGCTTATTTGGATATCAAGCGCTGATGATTAATAAAGC 480
Qy 497 TGATTTAAGGACTGGGGTAAGGAATGATGTGGCAGGTGGTACTATAGGCGT 556
Db 481 TGTCTTTCATGAGCTGGGGTAAGGAGTATGTAGCAGGTGTGTTACTATGCGCT 540
Qy 557 GACAAAGATTGACAGAAATGTTCTTCATTTACATTTGCTAATATTAACAGAACT 616
Db 541 GACAAAGATTGACAGAAATTTATTTCCATTTACATTTGCTAATATTAACAGAACT 600
Qy 617 GAAAAATAGCTTAAATAGCTTGGAGGATTTGAGCTGTTGATTAACATGCG 676
Db 601 AAAAATAGCTTAAATAGCTTGGAGGATTTGAGCTGTTGATTAACATGCG 660
Qy 677 CCAAGTACTGCTGTTGGTTGGAGCTGCTTCAAAAAAGCTATTAATTTATTC 736
Db 661 CCATTCAGTCTGTTGGGTTTGAAGCTCTTCAAAAAAGCTATTAATTTATTC 720
Db 737 TAGAAATGGGTG 749
Db 721 ACAAAATGGGTG 733

RESULT 4

AAx61505
ID AAX61505 standard; DNA; 540 BP.

AAx61505;

DT 19-JUL-1999 (first entry)

DE B. burgdorferi antigenic protein coding sequence, f933.nt.

KW Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

OS Borrelia burgdorferi.

PN WO9859071-A1.

PD 30-DEC-1998.

PF 18-JUN-1998; 98WO-US12718.

PR 03-SEP-1997; 97US-0057483.

PR 20-JUN-1997; 97US-0050359.

PR 22-JUL-1997; 97US-0053344.

PR 22-JUL-1997; 97US-0053377.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (MEDI-) MEDIMUNE INC.

PI Chol GH, Erwin AL, Hanson MS, Lathigra R;

DR WPI; 1999-189980/16.

DR P-PSDB; AAY19808.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop

PT products for the diagnosis, prevention and treatment of diseases

PT caused by Borrelia, particularly Lyme disease

XX Claim 1; Page 73; 275pp; English.

CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.

XX Sequence 540 BP; 145 A; 71 C; 123 G; 201 T; 0 other;

Query Match 53.8%; Score 402.8; DB 20; Length 540;
Best Local Similarity 85.6%; Pred. No. 3.1e-68;
Matches 462; Conservative 0; Mismatches 72; Indels 6; Gaps 1;

Qy 192 ATGATAGTATTTTAAATTTTATTTTGGTAACTCTTTGCTTTTCTAGTTTGCACAA 251
Db 1 ATGATTAATCTTTAAATTTTGTGTTTGGCACTTTTGTGTTTCTAGTTTGCACAA 60
Qy 252 GATGATCT-----AAAGACCTTTTAAATCTGGAGCGGAGAAAAATTTTGGTTAT 305
Db 61 GCTATGATTTCTAAATATGTCGCTTGGATGAGTGCTGAGAAAACTTTTGGTTAT 120
Qy 306 GAACTATTAAGAAAGATCTCTGTTACATTTTATTTGAACCTTTTGGTTGGG 365
Db 121 GAACTAGCAAGCAAGATCTTATTTACATTTTATTTGAACCTTTTGGTTGGG 180
Qy 366 ATAGTCTTTTGGTCAAGAGATATCTGAGGTTCTCTTATTTGATTTGATGCG 425
Db 181 ATAGGCTCTTTGCTCAAGGAGATTTCTTGAGGTTCTCTTATTTGATTTGATGCG 240
Qy 426 GTTGATATAGGTTAATATCTTACAGAGCTTATTTGACATCAAGATTTTGAATAT 485
Db 241 GTTGATATAGGCTTATATCTTACAGGCTTATTTGATCAAGAGCTTATGATAT 300
Qy 486 GCTAAAAAGCTGATTTTAACTGAGCTTGGGTAAGGAAATGATTTGGAGTGTGTT 545
Db 301 ACTAAAAAGCTGCTTTTCAATGAGCTTGGGTAAGGATTTGATGAGTGTGTT 360
Qy 546 ACTATGCTGTGACAGATTTGACAGAAATTTCTTCCATTTACATTTGCTAATATAT 605
Db 361 ACTATGCTGTGACAGATTTGACAGAAATTTCTTCCATTTACATTTGCTAATATAT 420
Qy 606 AACAGAGCTGAAAAATATGCTTAAATATATGCTTGGAGGATTTGAGCTTATTTGAT 665
Db 421 AATAGGAAGCTAAAAAATATGCTTAAATATATGCTTGGAGGATTTGAGCTTATTTGAT 480
Qy 666 ATTAACATGCGCCCAAGCTAGTGTCTTGGTTTGGAGCTGCTTTCAAAAAAGCTATTA 725
Db 481 GTTGCAATGGGCCAATCCAGCTCTTGGGTTTGAACCTCTTTCAAAAAAGCTATTTAA 540

RESULT 5

AAx61506
ID AAX61506 standard; DNA; 477 BP.

AAx61506;

DT 19-JUL-1999 (first entry)

DE B. burgdorferi antigenic protein coding sequence, f933.nt.

KW Antigenic protein; vaccine; Lyme disease; infection; detection; ss.

OS Borrelia burgdorferi.

PN WO9859071-A1.

PD 30-DEC-1998.

PF 18-JUN-1998; 98WO-US12718.

PR 03-SEP-1997; 97US-0057483.

PR 20-JUN-1997; 97US-0050359.

PR 22-JUL-1997; 97US-0053344.

PR 22-JUL-1997; 97US-0053377.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (MEDI-) MEDIMUNE INC.

PI Chol GH, Erwin AL, Hanson MS, Lathigra R;

DR WPI; 1999-189980/16.

DR P-PSDB; AAY19809.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by Borrelia, particularly Lyme disease
 PS
 XX
 XX Claim 1: Page 73; 275pp; English.

CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus.

XX Sequence 477 BP; 132 A; 62 C; 114 G; 169 T; 0 other;

Query Match 49.2%; Score 368.8; DB 20; Length 477;
 Best Local Similarity 85.9%; Pred. No. 9.7e-62;
 Matches 409; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

250 AAGATGATCTTAAAGCACTTTTAACTGGAGCGGAGAAAAATTTGGTTATGAAA 309
 1
 2 ATGATCTTAAATGATGGGTTGGATGAGTCTGAGAAAAACTTTGGTTATGAAA 61
 310 CTAATAGAAAGATCTCTTGACATTTTATGAACTTTTAAAGGCTCGGATAG 369
 1
 62 CTAAGCAAGCAAGATCTTATGACATTTTATGAACTTTTAAAGGCTCGGATAG 121
 370 GTTCTTTCTCAAGAGATCTCTGAGAGTTCTTATCTTGGATTGATGGGTTG 429
 1
 122 GCTCTTCTCAAGAGATCTCTGAGAGTTCTTATCTTGGATTGATGGGTTG 181
 430 GTATAGGTTAACTTACAGAGCTTATTTGACATCAAGATTTGATTAATATGCTA 489
 1
 182 GTATAGGCTTAACTTACAGAGCTTATTTGACATCAAGATTTGATTAATATGCTA 241
 490 AAAAAAGCTATTTAAGAGATCTGGGTAAGAGATGATGTTGGAGAGTGTGTTACTA 549
 1
 242 AAAAAAGCTATTTAAGAGATCTGGGTAAGAGATGATGTTGGAGAGTGTGTTACTA 301
 550 TGGCTGTGACAGATTTGACAGAAATGTTCTTCAATTTGATTAATATTAACA 609
 1
 302 TGGCTGTGACAGATTTGACAGAAATGTTCTTCAATTTGATTAATATTAACA 361
 610 GGAAGCTGAAAAATGCTTTAATATAGCTTGGAGAGATTTGAGCTAGTTTGAATATTA 669
 1
 362 GGAAGCTGAAAAATGCTTTAATATAGCTTGGAGAGATTTGAGCTAGTTTGAATATTA 421
 670 ACATGGGCAAGCTAGTGTCTGTTTGAAGCTGCTTCAAAAAAACCTATTA 725
 1
 422 CAATGGGCAAGCTAGTGTCTGTTTGAAGCTGCTTCAAAAAAACCTATTA 477

RESULT 6
 AAX20310/c
 ID AAX20310 standard; DNA; 1761 BP.

XX AAX20310;
 XX 04-MAY-1999 (first entry)

DE Borrelia burgdorferi polynucleotide sequence #63.

KM Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
 KM epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
 KM infection; diagnosis; characterisation; detection; ds.

OS Borrelia burgdorferi.

XX W09858943-A1.

XX 30-DEC-1998.

XX 18-JUN-1998; 98WO-US12764.
 XX 03-SEP-1997; 97US-0057483.
 XX 20-JUN-1997; 97US-0050359.
 XX 22-JUL-1997; 97US-0053344.
 XX 22-JUL-1997; 97US-0053377.

PA (HUMAN-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMUNE INC.

PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
 PI White OR;

DR WPI; 1999-081217/07.

PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the detection, diagnosis, characterisation, prevention
 PT and therapy of infections, particularly Lyme disease

PS Claim 1: Page 1045-1046; 1128pp; English.

AAX20248 to AAX20402 represent polynucleotide sequences isolated from
 CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for
 CC the detection, diagnosis, characterisation, prevention and therapy of
 CC Bb infections, e.g. Lyme disease. They can also be used for the
 CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
 CC to a family of motile, spiral-shaped bacteria called Spirochetes.
 CC Spirochetes are pathogenic in humans and Borrelia causes epidemic
 CC and endemic relapsing fever, and Lyme borreliosis, more commonly known as
 CC Lyme disease.

SQ Sequence 1761 BP; 548 A; 267 C; 268 G; 676 T; 2 other;

Query Match 13.1%; Score 98.4; DB 20; Length 1761;
 Best Local Similarity 49.7%; Pred. No. 5.2e-10;
 Matches 336; Conservative 0; Mismatches 331; Indels 9; Gaps 3;

1 TAGAATTTTCAACAAATTAAGATATTTGTTAAAGAAATGAAATTTGCTAATTTATGTTA 60
 1
 1031 TGGCATACAAAAAATAAGATAGCATTAATTTCAAAATTTAAAGAGTTTAAAGAACTTT 972
 61 AATCAGAGCTCTATTTGTAAGCGAATTTCCAGTACAAATTTGAAAGAAAGTTTA 120
 1
 971 ACATTAATAATCTGCTTAGAACAGCAAGAAATTTCAACATATCTTTAGAAAGATTA 912
 1
 121 AATAGTTCCAAAAGCTTTTAAATTTCAATTAATGCTACCATTAATACAGTTTAATA 180
 1
 911 ACTCCATTAATAATCTAATAATTTGCTACATTTGTTATCAAAACTTAATTTAG 852
 1
 181 AAGG--GGTTTATGAATTAAGTTTAAATTTTAAATTTTGAATCTTTTGTGCTTTT 237
 1
 851 GAGGTAATTAATTAAGAAAAAATTTTCAATTAATTAATTTTAAATTTTAAATTAAC 792
 1
 238 CTAGTTTGTCTCAAGATGATTTCAAAAGCACTTTTAAATCTGGAGCGGAGAAAAATTTT 297
 1
 791 AATCTTTGTCAACAGATTAACCTTGAAGAGTGTGAGT-GGAGATATTCGACCTGTA 733
 1
 298 TGGTTATGAATAATTAAGAAAGATTTCTCTGTACCATTTTAAATTAAGCTTTTATG 357
 1
 732 TGAATATGAAGCAAGAAAGCAACATCTAGACCATTTACTTTTGAATATCTTTTGTG 673
 1
 358 GGTTCGGATAGGTTCTTTTGTCTCAAGAGATATCTTGGAGGTTCTTATTTGTGAT 417
 1
 672 CTTTAAAGGATAGATCTTTTGTCAAGAGATTAATTTGTTGTTGCTGCACTTGTGCT 613
 1
 418 TTGATCGGTTGTTAAGGTTAATCTTACAGAGCTTATTTGATCAATCAAGAGATTTG 477
 1
 612 CTCAGTGTCTTGAAGG-----AATCTTATTAATGCTGACACATGATCGTAGAGGTA 558
 1
 478 ATAATATGCTAATAAAGCTGATTTTAAGTGTGAGGTTAAGGAAAGATTTGCGAG 537
 1
 557 TTGAGCGGTTTACAGAAACACAGCAACAGTATTAATCCGAGAGAGTATTAATCAGGAATAG 498


```

XX (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
XX White OR;
XX WPI: 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of infections, particularly Lyme disease
XX
XX Claim 1: Page 951-955; 1128pp; English.
XX
XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for
XX the detection, diagnosis, characterisation, prevention and therapy of
XX Bb infections, e.g. Lyme disease. They can also be used for the
XX production of biosynthetic products, e.g. enzymes. Borrelia belongs
XX to a family of motile, spiral-shaped bacteria called Spirochetes.
XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and
XX endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX Lyme disease.
XX
XX Sequence 7074 BP: 3064 A; 906 C; 866 G; 2233 T; 5 other;
XX
XX Query Match 11.0%; Score 82.2; DB 20; Length 7074;
XX Best Local Similarity 49.9%; Pred. No. 7.2e-07;
XX Matches 333; Conservative 0; Mismatches 308; Indels 26; Gaps 4;
XX
XX 13 CAAATTAAGATATGTTTAAGAATGGAATGCTAATTTTAAAGTTAAATCAAGAGCT 72
XX 2982 CTAATAGAGAGCTCTATGATATCATCAAAAAGTATTAATAATCAAAATTAACCTAT 3041
XX 73 CTAATGTAAGGCAATTCGAGTAACATTTGAAAAAGTTAAATTAATAGT-TCGAA 131
XX 3042 TTATTAACAACCTCATTAACCCAAATTCATAGAAATTAATAGTTAATGTTCTTGG 3101
XX 132 AAGCCTTTTAAATTCATTAAATGCTACCA-TAATACAGTTTAATTAAGGGGTTT 190
XX 3102 GGGTTTATTAATTAATGCTAATTAATTAATTAATTAATTAATTAATTAATTAATCA 3161
XX 191 TATGATTAAGTTTAAATTTTAAATTTGTAATCTTTGCTTTTCTAGTTTCTCTCA 250
XX 3162 TATGAAAAAATTTTAACATGATTAATTTTAACTTAACAAATCAATCTTCAAC 3221
XX 251 AGATGATTTCTAAAGCACTTTTAATCTGGAGCGGAGAAAAATTTTGGTTATGAAC 310
XX 3222 ACAAG----TAAAGCTGAAAAAAGTGTGGAATATTGAACCATTAATGAATATAAAG 3278
XX 311 TAAATAAAGAAATTCCTCTGTACCATTTTAAATGAACCTTTTAAAGGTTGGGATAG 370
XX 3279 CCAAAAAGCACTATCTAGCAACATTCCTTTGATTAATTTTAACTTTAAGATAGG 3338
XX 371 TTTCTTGTCAAGGATATCTGTGAGGTTCTTATCTTATCTTGTGATGAGCGGTTGG 430
XX 3339 ATCTTTGTCCAAAGAGATTAATTTGGTGGCGCAGCTGGGATCTCAGTTATTATGA 3398
XX 431 TATAGGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 490
XX 3399 AAGGAATACCTTGATAGCTGGAATATTTCTGGCCATACAGATGAGAAACAAGACAAC 3458
XX 491 AAAAGCTGATTTTAACTGACCTTGGGGTAAAGGAATGATGTTGGCAGGTGGTTACTAT 550
XX 3459 AACTGGGCAATAT-----AATTAACAAGATAGGATGGAGCAGAT 3497
XX 551 GGCTGTGAACAATGTAAGAAATTTGTTCCATTTAATTAATTAATTAATTAATTAACAG 610
XX 3498 TATAGCATCCCAATAGCTCACTATTAATTTCCATTTAATTTGCAAAATTAACACATGC 3557
XX 611 GAAGTGAAAAATAGTCTTAATATAGCCTTGGAGGATTTGAGCCAGTTTGTATTTAA 670

```

```

DB 3558 AATCTTAATAAAGAGAGCTGGCATGATATGCGGGTTTGAACCAATTTGATATGG 3617
XX
XX 671 CATGGC 677
XX
DB 3618 AATTAAGC 3624
XX
XX
XX RESULT 9
XX AAS46489
XX ID AAS46489 standard; DNA; 7110 BP.
XX
XX AAS46489;
XX
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #211.
XX
XX Human: tumour suppressor gene; oncogene; antitumour; cytostatic;
XX cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
XX
XX Homo sapiens.
XX
XX WO200168912-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP02955.
XX
XX 15-MAR-2000; 2000DE-1013847.
XX
XX 06-APR-2000; 2000DE-1019058.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumour suppressor
XX genes and oncogenes, useful in designing primers and probes for
XX analysing diseases associated with cytosine methylation state e.g.
XX cancer
XX
XX Claim 1: SEQ ID No 211; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and
XX oncogenes having a sequence taken from 536 (actually 533 since
XX numbers 408, 458 and 500 are missing from the sequence listing) sequences
XX (Ss) and sequences complementary to (Ss). The nucleic acid may be a
XX peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
XX form part of a set of probes for detecting the cytosine methylation state
XX and/or single nucleotide polymorphisms and also to be used in an
XX array for analysing diseases associated with CpG dinucleotides e.g.
XX cancers and tumours. The probes can also be used in a method for
XX ascertaining genetic and/or epigenetic parameters for the diagnosis
XX and/or therapy of existing diseases or the predisposition to specific
XX diseases, by analysing cytosine methylations. The parameters may be
XX compared to another set of genetic and/or epigenetic parameters, the
XX differences serving as basis for diagnosis and/or prognosis events which
XX are disadvantageous to patients. The present sequence is one of the
XX 533 genomic sequences derived from tumour suppressor genes and
XX oncogenes.
XX
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from Wipo at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 7110 BP: 1749 A; 289 C; 1822 G; 3250 T; 0 other;

```

Query Match 8.8%; Score 66; DB 22; Length 7110;
Best Local Similarity 51.0%; Pred. No. 0.0009;
Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 35 AATTGAATTCGTAATTTATGTTAATCAAGAGCTCATTTGTGTAGGCAATTCGAG 94
DB 1672 AATTGAATTTTAAAGATATTTATTTTGAATTTTGTAGATGGGTTAGTTA 1731
QY 95 TAACAATTTGAAAAAGTTAATTTAATAGTTCACAAAGCCTTTTAAATTCATTAA 154
DB 1732 TTGGGTAGAGAAAGAAAGAAATTAATAAAGATGATATAGATTTTATTTT 1791
QY 155 TATGCTACCATATATACAGTTTATAAAGGGTTTATGATAGATTTTAAATTTTAT 214
DB 1792 AGAATATATTTAATATATAGAAATAGAGAGATTTATTTATTAATTTTAAATATAT 1851
QY 215 TTGTGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATTTCTAAAGCCTTTTAA 274
DB 1852 TTGATGTTGTTGGTGGTGGTTGTTAATTTGTTAATATTTGATTTTATTTTAAATTA 1911
QY 275 TCTGGAGCGGAGAAAAATTTTGTATTAGAACTAATATAGAAAGATTCCTGTAC 334
DB 1912 TTAGAGAGAAATAGAGGATTTGATGTTTAGAGAGTAGTAATAATTTATTTTATTTT 1971
QY 335 ATTTT 340
DB 1972 TTTT 1977

RESULT 10
ABL70587
ID ABL70587 standard; DNA: 7110 BP.
AC ABL70587;
XX
DT 01-JUL-2002 (first entry)
XX
DE Chemically treated cell signalling DNA sequence#239.
XX
KW Cell signalling; cytosine methylation; cell signalling disease;
KM cancer; tumour; cytostatic; ds.
XX
OS Unidentified.
XX
PN WO200202807-A2.
XX
10-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-EP07471.
XX
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-154758/20.
DR
XX
XX Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling -
XX
XX
PS Claim 1; SEQ ID NO 477; 24pp+sequence listing; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is

CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;

Query Match 8.8%; Score 66; DB 24; Length 7110;
Best Local Similarity 51.0%; Pred. No. 0.0009;
Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 35 AATTGAATTCGTAATTTATGTTAATCAAGAGCTCATTTGTGTAGGCAATTCGAG 94
DB 1672 AATTGAATTTTAAAGATATTTATTTTGAATTTTGTAGATGGGTTAGTTA 1731
QY 95 TAACAATTTGAAAAAGTTAATTTAATAGTTCACAAAGCCTTTTAAATTCATTAA 154
DB 1732 TTGGGTAGAGAAAGAAAGAAATTAATAAAGATGATATAGATTTTATTTT 1791
QY 155 TATGCTACCATATATACAGTTTATAAAGGGTTTATGATAGATTTTAAATTTTAT 214
DB 1792 AGAATATATTTAATATATAGAAATAGAGAGATTTTATTTAATTTAATATATAT 1851
QY 215 TTGTGTAATCTTTTGTCTTTTCTAGTTTGTCTCAAGATGATTTCTAAAGCCTTTTAA 274
DB 1852 TTGATGTTGTTGGTGGTGGTTGTTAATTTGTTAATTTGATATTTTATTTTAAATTA 1911
QY 275 TCTGGAGCGGAGAAAAATTTTGTATTAGAACTAATATAGAAAGATTCCTGTAC 334
DB 1912 TTAGAGAGAAATAGAGGATTTGATGTTTAGAGAGTAGTAATAATTTATTTTATTTT 1971
QY 335 ATTTT 340
DB 1972 TTTT 1977

RESULT 11
ABL33556
ID ABL33556 standard; DNA: 7110 BP.
AC ABL33556;
XX
XX 26-MAR-2002 (first entry)
DT
XX
DE Human immune system associated gene SEQ ID NO: 1529.
XX
XX Human immune system associated gene SEQ ID NO: 1529.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytostatic; noctropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200200928-A2.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP07537.
PF
XX
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX

PI Olek A, Plepenbrock C, Berlin K;
 XX WPI: 2002-130909/17.
 XX
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 XX Claim 1: SEQ ID NO 1529; 32pp + Sequence Listing; German.
 XX
 XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 XX Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;
 SQ
 Query Match 8.8%; Score 66; DB 24; Length 7110;
 Best Local Similarity 51.0%; Pred. No. 0.0009;
 Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
 OY 35 AATTGAATTCCTAATTTATGCTTAATCAAGACCTATTTGGTAACGATTTCCGAG 94
 Db 1672 AATTGAATTTTAAAAAGATTTATTTTGAATTTTCTAGTTAGATGAGGTTAGTTA 1731
 OY 95 TAACATTTGAAAAAGTTAAATTAATGATTCACAAAGCCTTTTAAATTTCAATTA 154
 Db 1732 TTGGGCTAGAAAAGAAAAGAAAATTTGAAAAAGATGTTATTTTATTTT 1791
 OY 155 TATGCTACCAATATACCACTTAATTAAGGGCTTTTATGAATTAAGTTTAAATTTTAT 214
 Db 1792 AGAATTAATTTAATTAATAGATAGAGAGATTTTATTTAATTTTATTAATATAT 1851
 OY 215 TTGGTAATCTTTTGTGCTTTTCTAGTTTCTCAAGATGATTCAAAAGCACTTTAA 274
 Db 1852 TTGATGTTGTTTGTGGTGGTTAAATATTTTAAATTTGATTTTATTTTATTA 1911
 OY 275 TCTGGAGCGGGAAGAAAATTTTGGTTATGAACCTAATTAAGAAATCTCTTAC 334
 Db 1912 TTAGAGAGAGTAGAGGAGATTGATGTTTAGAGAACTAATAATTTATTTTATTTT 1971
 OY 335 ATTTT 340
 Db 1972 TTTT 1977
 ASUFL 12
 AAS61290
 ID AAS61290 standard; DNA; 7110 BP.
 XX
 XX AAS61290;
 AC
 XX 29-JAN-2002 (first entry)
 DT
 XX
 XX Human gene regulation-associated gene oligonucleotide #245.
 DE
 XX
 XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KM cardiac damage; inflammatory response; Hemophilia; Werner syndrome;
 KM asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KM renal disease; Preclampsia; cardiac allograft vascular disease;
 KM colorectal cancer; thyroid cancer; oesophageal cancer; ds. tumour;
 KM immunostimulant; cardiac; antiinflammatory; coagulant; antisthmatic;
 KM nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200177375-A2.
 PN
 XX
 XX 18-OCT-2001.

XX
 PF 06-APR-2001; 2001WO-EP03968.
 XX
 XX 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX
 XX Olek A, Plepenbrock C, Berlin K;
 PI
 XX WPI: 2002-017470/02.
 DR
 XX
 XX New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease
 XX
 XX Disclosure: SEQ ID NO 251; 26pp; English.
 PS
 XX
 XX The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 CC preclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 SQ
 Sequence 7110 BP; 1749 A; 289 C; 1822 G; 3250 T; 0 other;
 Query Match 8.8%; Score 66; DB 24; Length 7110;
 Best Local Similarity 51.0%; Pred. No. 0.0009;
 Matches 156; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
 OY 35 AATTGAATTCCTAATTTATGCTTAATCAAGACCTATTTGGTAACGATTTCCGAG 94
 Db 1672 AATTGAATTTTAAAAAGATTTATTTTGAATTTTCTAGTTAGATGAGGTTAGTTA 1731
 OY 95 TAACATTTGAAAAAGTTAAATTAATGATTCACAAAGCCTTTTAAATTTCAATTA 154
 Db 1732 TTGGGCTAGAAAAGAAAAGAAAATTTGAAAAAGATGTTATTTTATTTTATTT 1791
 OY 155 TATGCTACCAATATACCACTTAATTAAGGGCTTTTATGAATTAAGTTTAAATTTTAT 214
 Db 1792 AGAATTAATTTAATTAATAGATAGAGAGATTTTATTTAATTTTATTAATATAT 1851
 OY 215 TTGGTAATCTTTTGTGCTTTTCTAGTTTCTCAAGATGATTCAAAAGCACTTTAA 274
 Db 1852 TTGATGTTTGTGGTGGTTAAATATTTTAAATTTGATTTTATTTTATTA 1911
 OY 275 TCTGGAGCGGGAAGAAAATTTTGGTTATGAACCTAATTAAGAAATCTCTTAC 334
 Db 1912 TTAGAGAGAGTAGAGGAGATTGATGTTTAGAGAACTAATAATTTATTTTATTTT 1971
 OY 335 ATTTT 340
 Db 1972 TTTT 1977


```

Db 6669 TTTTAAAGAGTTATTTCGTTTTATTATAGAGTAGATTTTATTTTTAAAGTACTTT 6728
Qy 374 TTTTGCTCAAGAGATATCCTTGAGGCTCTCTTATTCTTGATTTGATCGGTGTAT 433
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6729 TATAGTTTAAAGTTATATTTTTTTTTTTTATTAGTATATTTGTTAAATAA 6788
Qy 434 AGGTTAATACCTTACAGAGCCTTATTGGACATCAAGATTTGATATATATGCTAAAA 493
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6789 TTTTTTTTAAATATATAGTTTTTTTGTAAAGGTAGATGAATTTATTTAGT 6848
Qy 494 AGCTGATTTTAACTGACT 512
    || || || | | | | | | | | | | | | | | | | | | | | | |
Db 6849 TTAGATTTTAGTTTAT 6867
    || || || | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: March 27, 2003, 01:32:55
 Job time : 204.602 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 27, 2003, 01:16:47 ; Search time 1051.13 Seconds
(without alignments)
11540.380 Million cell updates/sec

Title: US-09-508-487-22

Perfect score: 749

Sequence: 1 tagaatttcacacaaataa.....attatctagaataatgggtg 749

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estipl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74.8	10.0	1101	17	CNS00LTX2
2	72.4	9.7	1101	17	CNS00BO7
3	71.8	9.6	1101	17	CNS0039G
4	69.6	9.3	996	17	CNS00PFU
5	68.2	9.1	1101	17	CNS0100X
6	68	9.1	1225	17	CNS0161D

7	67.8	9.1	1013	17	CNS06RPQ	AL412260 T7 end of
8	66.6	8.9	910	17	CNS0168P	AL442826 Anopheles
9	65	8.7	1101	17	CNS00BEVL	AL069706 Drosophila
10	64.8	8.7	1201	17	CNS016EQ	AL106652 Drosophila
11	64.2	8.6	928	17	CNS00DKY	AL071865 Drosophila
12	64.2	8.6	1092	17	CNS020K7	AL175696 Tetradon
13	62.8	8.4	1038	17	CNS01TEL7	AL166600 Tetradon
14	62.4	8.3	1204	17	CNS016E2	AL106628 Drosophila
15	62	8.3	576	17	CNS035N7	AL228940 Tetradon
16	62	8.3	893	17	BH188563	BH188563 036_O_12-
17	62	8.3	893	17	CNS075JW	AL625502 T3 end of
18	61.6	8.2	1101	17	CNS00LT2	AL078714 Drosophila
19	61.4	8.2	893	17	CNS013XE	AL103436 Drosophila
20	61.4	8.2	1101	17	CNS004ZM	AL055440 Drosophila
21	61.4	8.2	1303	14	BH924677	BH924677 AGENCOURT
22	61.2	8.2	1001	17	CNS0155H	AL105023 Drosophila
23	61.2	8.1	1101	17	CNS002DB	AL062360 Drosophila
24	61	8.1	1101	17	CNS00EVL	AL069706 Drosophila
25	61	8.1	1101	17	CNS012I9	AL101595 Drosophila
26	60.4	8.1	865	12	BF664686	BF664686 HV-CEA001
27	60.2	8.0	1101	17	CNS00DT7	AL075293 Drosophila
28	60	8.0	1300	10	BE420576	BE420576 HMO000.D0
29	59.8	8.0	524	17	CNS01U90	AL167541 Tetradon
30	59.8	8.0	1184	17	CNS04P4P	AL300850 Tetradon
31	59.6	8.0	987	17	CNS014P0	AL104456 Drosophila
32	59.6	8.0	994	17	CNS04187	AL169872 Tetradon
33	59.6	8.0	1101	17	CNS017KE	AL108152 Drosophila
34	59.4	7.9	865	17	BH180441	BH180441 016_H_20-
35	59.4	7.9	865	17	CNS007M	AL161793 T3 end of
36	59.4	7.9	1101	17	CNS003BD	AL064091 Drosophila
37	59.2	7.9	928	17	CNS00DKY	AL071865 Drosophila
38	59.2	7.9	1201	17	CNS016FX	AL106695 Drosophila
39	59	7.9	734	17	CNS010MP	AL099163 Drosophila
40	59	7.9	876	17	CNS009G1	AL053529 Drosophila
41	59	7.9	1092	17	CNS020K7	AL175696 Tetradon
42	58.9	7.9	1101	17	CNS016DT	AL106619 Drosophila
43	58.8	7.9	1101	17	CNS017B3	AL107817 Drosophila
44	58.6	7.8	876	17	CNS009G1	AL053529 Drosophila
45	58.6	7.8	914	17	CNS002JY	AL097768 Drosophila

ALIGNMENTS

RESULT 1 1101 bp DNA linear GSS 14-JUN-1999
CNS00LTX2
LOCUS
DEFINITION
BACR48P19 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL078714.1 GI:5102004
VERSION
KEYWORDS
SOURCE
ORGANISM

Drosophila melanogaster.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

REFERENCE
GENOSCOPE.
TITLE
JOURNAL

COMMENT

Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPC1-98 and was constructed by partial

ECORI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source

```
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR48P19"
/clone_lib="RPC1-98"
/note="end : TET3"

BASE COUNT      469 a      6 c      69 g      151 t      406 others
ORIGIN
```

Query Match 10.0%; Score 74.8; DB 17; Length 1101;
Best Local Similarity 24.5%; Pred. No. 0.00074;
Matches 172; Conservative 203; Mismatches 322; Indels 5; Gaps 1;

```
11 AACAAATAAGATATTTGTTAAAGATTGAATTCCTAATTTTATGTTAAATCAAGAAG 70
   ||||| ||| | | ||||| ||| | | ||||| ||| | | ||||| ||| | |
Db 381 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 440

QY 71 CTATATGCTAAGCGAATTTGAGTACAAATTTGAAAAAGTTAAATTAATGCTCA 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 441 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTTCTTTT 500

QY 131 AAGCCTTTTAAATTCATTATATGCTACCAATACCGTTTAAATGAAGGGTTT 190
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 501 ATTTTWTTTTADADKMAAMDTTTTTAMTTWTWTTAAATTTTCTTTTCTTTT 560

QY 191 TATGAATAAGTTTAAATTTTATTTGTAATCTTTGCTTTGCTTTGCTGCTCA 250
   ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 561 DMTKAKTAATAAAMWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 620

QY 251 AGATGATTTCTAAAGCACTTTTAAATCTGGAGCGGAGAAAAATTTTGGTTATGAAC 310
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 621 TTTTWTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 680

QY 311 TAATAG-----AAGATTCCTCTGTACCAATTTTATGAACCTTTTAAAGGTTGCG 365
   || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 AAAAAWTTTDTWAAAMWTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 740

QY 366 ATAGCTCTTTTGTCTCAAGAGATATCCTTGGAGGCTCTCTATCTTTGATTTGATCG 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 741 DKDKKGGGKKKKGGGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 800

QY 426 GTTGCTAATAGGTTAACTTAACAGAGCTTATTTGACATCAAGATTTGATTAAT 485
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 801 TTTTGGGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 860

QY 486 GCTAAAAAGCTGATTTAAGTGACCTGGGTAAGGAGATGATGTTGGCAGTGCTG 545
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 861 KKKKKKKGGGKKKDAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 920

QY 546 ACTATGCTGTGACAGATTTGACAGAAATTTGTTCCATTTACATTTGCTAATATAT 605
   : : | : : : : : : : : : : : : : : : : : : : : : : : :
Db 921 KKKGGGGGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 980

QY 606 AACAGGAAGCTGAAAAATAGTCTTAATATAGCTTTGGAGATTTGAGCCTAGTTT 665
   : : | : : : : : : : : : : : : : : : : : : : : : : : :
Db 981 TDAATKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1040

QY 666 ATTAACATGGCCAGCTAGTCTCTTGTTTGGAGCTGCT 707
   | | : : : : : : : : : : : : : : : : : : : : : : : :
Db 1041 KTTGGKDKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 1082
```

RESULT 2
CNS00E07
LOCUS CNS00E07 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION *Drosophila melanogaster* genome survey sequence TET3 end of BAC:

BACR29P01 of RPC1-98 library from *Drosophila melanogaster* (fruit fly), genomic survey sequence.

AL069440

VERSION AL069440.1 GI:4949583

KEYWORDS GSS.

SOURCE

ORGANISM

Drosophila melanogaster.

Drosophila melanogaster.

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; *Drosophila*.

1 (bases 1 to 1101)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley *Drosophila* Genome Project (BDGP).
The BDGP is constructing a physical map of the *Drosophila*
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP *Drosophila*
melanogaster BAC library was prepared by Kazuhiro Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPC1-98 and was constructed by partial
ECORI digestion of *Drosophila* DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers

FEATURES

```
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR29P01"
/clone_lib="RPC1-98"
/note="end : TET3"
```

```
BASE COUNT      366 a      66 c      104 g      351 t      214 others
ORIGIN
```

Query Match 9.7%; Score 72.4; DB 17; Length 1101;
Best Local Similarity 36.2%; Pred. No. 0.0019;
Matches 146; Conservative 67; Mismatches 190; Indels 0; Gaps 0;

```
11 AACAAATAAGATATTTGTTAAAGATTGAATTCCTAATTTTATGTTAAATCAAGAAG 70
   || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 699 AATTAAAAAAMAAWTTTAAATTAAMAAATTAATTTTATTTTAAATTAAMAAWTTAA 758

QY 71 CTATATGCTAAGCGAATTTGAGTACAAATTTGAAAAAGTTAAATTAATAGTTCCA 130
   || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 759 AATTATWATWAAAMWTTTAAAMAAATTTTAAATTTTAAATTTTAAATTTTAAWTT 818

QY 131 AAGCCTTTTAAATTCATTATATGCTACCAATACCAAGTTTAAATGAAGGGTTT 190
   | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 819 TAAATATATATATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAWTTG 878

QY 191 TATGAATAAGTTTAAATTTTATTTTGTATATCTTTGCTTTTCTGTTTCTGCTCA 250
   || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 879 TAAATWAAATTTAAATTTTAAAMAAATGTTTAAATTTTAAATTTTAAATTTG 938

QY 251 AGATGATTTCTAAAGCACTTTTAAATCTGGAGCGGAGAAAAATTTTGTATGAAC 310
   | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 939 TATATAAAAAAMWATTATGATTAAMAAATTTTAAATTTTAAATTTTAAATTTTAA 998

QY 311 TAATAAGAAAGATTTCTGTTGACATTTTAAATTTTAAATTTTAAATTTTAAATTT 370
   || | | | | | | | | | | | | | | | | | | | | | | | | |
Db 999 AAAAAATATWTTAAAMAAATTTGTTTAAAMAAATTTTAAATTTTAAATTTTAA 1058

QY 371 TTCTTTGCTCAAGAGATTCCTGAGGTTCTCTTATCTCT 413
   | : : : : | | | | | | | | | | | | | | | | | | | | |
Db 1059 TGGGDAARAAARAKAKGCTTTGTTKTAATTTTCTTTTCTTTT 1101
```

[illegible]

Qy	361	TCGGATGATGGTCTTTTGCTCAGAGATATCCCTTGAGGTTCTCTATTCCTGGATTG	420
Dy	740	AGTAAKRRKRRKRRKRRKRRDADDDADTDARDDRKRGGADAGCKTKRRRRRD	681
Qy	421	ATGCGGTGGATGATGAGCTATACATACAGAGCTATTGTGACATCAAGATTGGATA	480
Dy	680	RATMDRTDAMWADAMWTTTDTDDTDMKRDORRRKGARRRRRTTARAAMDWTKAMDMA	621
Qy	481	ATAATGCTAAAAAGACGTATTTAACTGGACCTTGGGGTAAGGAATGATGTTGGCAGCTG	540
Dy	620	KWMDKTRADFRMDRNAADTWTARAKADRWMAKARARARARBRARARADRRWTKGTTTA	561
Qy	541	TGGTTACATAGCGCTGACAGATTTGACAGAAATGTTCTTCATTTTCATTTGCTATA	600
Dy	560	TWTWTAATAAAMWAMWAMWATTTATWTTTWTWTTTWTWTTTWTWTTTAAWMAWWTW	501
Qy	601	ATTATTAACGAGAGAGCTGAAAAATAGTCTTAATATAGCCCTGGAGATTTGAGCCTACTT	660
Dy	500	AAATTAATAAATAAATAAATTTTTTTTTTTTTTAAWMTAAWMTWTTTWTWTAATTT	441
Qy	661	TTGATATTACAT 673	
Dy	440	TTTWTWTTWMAAT 428	
RESULT 4			
CNS00F0H/c			
LOCUS	CNS00F0H 996 bp DNA linear GSS 03-JUN-1999		
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR31021 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL071063		
VERSION	AL071063.1 GI:4951105		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster.		
ORGANISM	Drosophila melanogaster.		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 996)		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .		
FEATURES	Location/Qualifiers		
source	1..996		
	/organism="Drosophila melanogaster"		
	/db_xref="taxon:7227"		
	/clone="BACR31021"		
	/clone_id="RPc1-98"		
	/note="end : TET3"		
BASE COUNT	383 a 164 c 81 g 171 t 197 others		
ORIGIN			
Query Match	9.3%	Score 69.6;	DB 17; Length 996;
Best Local Similarity	36.2%;	Pred. NO. 0.0061;	

Matches	164	Conservative	65	Mismatches	224	Indels	0	Gaps	0
OY	1	TAGAAATTTTCAACAATATAGATATFTGTTAAAGACATTGGAATTCCTAATTTATGCTTA	60						
Db	892	WMMWATTTTTTWTWMAATTTAARTTWRATTTWMAWMTAMWTTTTTTTTTTTTTTTTTA	833						
OY	61	AATCAAGAAGCTCATATGTTGAACGAATTTGAGTAACAAATTTGAAAAAAGTTAAATTA	120						
Db	832	ATWDTTWTTTTTTTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	773						
OY	121	AATAGTTCACAAAGCCCTTTTAAATTTTCATTAAATATGCTACCAATATACAGTTAATA	180						
Db	772	TTTTTTTTTTTTTTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	713						
OY	181	AAGGCGTTTATATGAAATAGTTTAAATTTTATTTATTTTGGTAATCTTTTGCTTTTCTA	240						
Db	712	WTTTTTTTTTTTTTTTTTWTWMAATTTTTTWTWMAATTTATTTATTTATTTATTTATTT	653						
OY	241	GTTTGTCTCAAGATGATTTCTTAAGACACCTTTTAAATCTGAGAGCGGAGAAAAATTTTG	300						
Db	652	TTTTTTTTRRMTWTATATWATATTTWATTTWATTTAKAMTTWKAATTTTGGAAAMWTTT	593						
OY	301	TTTATGAACCTAATATAGAAAGATTCTCTGTACCACTTTTATATGAACTTTTATAGGCT	360						
Db	592	TTWMAATATATKADTAMWMTTWTATTTTTTTTTTATTTATTTATTTATTTWKAATTTTT	533						
OY	361	TCGGATATAGGTTCTTTTCTCTCAAGAGATATCCCTGAGAGGTTCTCTATTCCTGATTTG	420						
Db	532	TTTTTTTTTATTTTTTTTTTCCGTTTTTTTTTTTTTTTTTTTTTTTTTCTRKTTTTTTNN	473						
OY	421	ATCGCGTTGGTATAGGTTAATATCTTACAGAGAG	453						
Db	472	NNDKGRGGGGGGGCVVVVMMNMAAAAAAGGG	440						

RESULT 5

CNS0100X

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

BASE COUNT

ORIGIN

CNS0100X

Drosophila melanogaster genome survey sequence Sp6 end of BAC BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL098379

AL098379.1 GI:5609990

GSS.

Drosophila melanogaster.

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

Location/Qualifiers

1..1101

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="BACN03G04"

/clone_lib="DrosBAC"

/plasmid="pBelobAC11"

/note="End : Sp6"

195 a 108 c 131 g 161 t 506 others

[illegible]

FEATURES	Location/Qualifiers
source	1. .1225
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone="BACN15C18"
	/clone="1b-DrosBAC"
	/plasmid="pBelOBAC11"
	/note="end : SP6"
BASE COUNT	266 a 128 c 38 g 368 t 425 others
ORIGIN	
Query Match	9.1%, Score 68; DB 17; Length 1225;
Best Local Similarity	33.4%, Pred. No. 0.011;
Matches 153; Conservative 94; Mismatches 203; Indels 8; Gaps 1;	
OY	4 AATTTCACAACAATTAAGATATGTTTAAAGATGGAATTCCTAATTTATGTTAAAT 63
DB	1225 AMWTWATWATATATTTTATKATTTTWWANAMAMMMWAMAMAAAMWTWTTATTAANA 1166
OY	64 CAAGAAGCTATATGTAAGCGAATTCGAGTACAAATTTGAAAAAGTTAAATTTAAAT 123
DB	1165 TATATWATWATATATTTTWWAMAAAMWTTTTATATWATWATWATWATWATWATW 1106
OY	124 AGTTCCAAAGCGTTTAAATTTTCATTAATATGCTACATATATGCAAGTTAAAG 183
DB	1105 WAMWAMWAMWAMWAMWAMWATWWTWKAAT-----ATWAAATWATWATWATWATW 1054
OY	184 GGGTTTATATGCAATGATTTTAAATTTTATTTTGTGATATCTTTGCTTTTCTAGT 243
DB	1053 ATWAMAMWTTWAMAAATATATTTTTTTWTWTTWTTWAMAAAAAMAAATATATW 994
OY	244 TTGCTCAAGATGATTTCTAAAGCACTTTTAAATCTGGAGCGGAGAAAAATTTGCTT 303
DB	993 WATAAAMWTTAAAAAAATTAATWTTATTTTAAATWTTTTRAAAMWTTTTTTWT 934
OY	304 ATGAACATTAATGAAGAAATCTCTGTGACCACTTTTATGAACTTTTATAGGCTT 363
DB	933 TTAATAAATWTTTATWTKTTWTRRRRRRAATGATWTTTTTTTTTTTTTATATATTT 874
OY	364 GGATAGCTTTCTTCTGCAAGAGATATCTTGTGAGCTTCTGTATGCTGATTTGATG 423
DB	873 TTTTWTNHTTWTGTTNN 814
OY	424 CGGTGCTATAGCGTTAATCTACTACAGACCTTATTTG 461
DB	813 GAGATTTTAAATTTTNNRTDWTARAANAAMDDTTWNTK 776
LOCUS	1013 bp DNA linear GSS 05-JUL-2001
DEFINITION	T7 end of clone AM0A016E05 of library AM0A from strain CL1B 89 of
ACCESSION	AL412260
VERSION	ALA12260.1 GI:12182622
KEYWORDS	GSS.
SOURCE	Yarrowia lipolytica.
ORGANISM	Yarrowia lipolytica
REFERENCE	Saccharomycetales; Dipodasaceae; Yarrowia.
AUTHORS	Soucié,J.L., Aigle,M., Attigienave,F., Blandin,G., Biolotin-Pukhara,M., Bon,E., Broctier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durren,P., Leplingle,A., Liorente,B., Malpertuy,A., Neuveglise,C., Ozler-Kalogeropoulos,O., Potier,S., Saurin,M., Tekala,F., Toffano-Nicohe,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE	20584711
PUBMED	11152876

REFERENCE	2 (bases 1 to 1013)
AUTHORS	Casaregola,S., Neuvéglise,C., Lepingle,A., Bon,E., Feynerol,C., Artiguenave,F., Wincker,P. and Gaillardin,C.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia lipolytica
JOURNAL	FEBS Lett. 487 (1), 95-100 (2000)
MEDLINE	20584727
PUBMED	11152892
REFERENCE	3 (bases 1 to 1013)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers
SOURCE	1..1013 /organism="Yarrowia lipolytica" /strain="CLIB 89" /db_xref="taxon:4952" /clone="AM0AA015R05" /clone_lib="AM0AA" /note="end : 77"
BASE COUNT	401 a 39 c 35 g 450 t 88 others
ORIGIN	
Query Match	9.1% Score 67.8; DB 17; Length 1013;
Best Local Similarity	43.4%; Preg. No. 0.013; Mismatches 202; Indels 0; Gaps 0;
Matches 171; Conservative	21; Mismatches 202; Indels 0; Gaps 0;
Dy	26 TGTAAAGATTGCAATTTTATGCTTAATCAAGACCTCATTTGTAACG 85
Dy	541 WNTTATAAATATATMTTATATATNATATAAATAATATATTAATTAATAAAA 600
Oy	86 AATTTCGAGTACAATTTGAAAAAGTTAAATTTAAATAGTCCAAAAGCCTTTTAA 145
Dy	601 AMATTWTTTNTTATTAATTAATTTTAAAAATTTTAAATATATTTTWMNAT 660
Oy	146 TTTCATTAATTCACATATACCAGTATTAAGAAGGGTTTTATGAATAACTTTT 205
Dy	661 AAATTTATAAATTAATAATTTTATTAATTAATMAATATATATAAATTTWTAA 720
Oy	206 AATTTTATTTTGTAATCTTTTGCTTTTCTAGTTTGCTCAAGATGATCTTAAAG 265
Dy	721 AAAATTTATATATATTAATTTATTTTWWTTTAAAAAATANTATTTWMAAT 780
Oy	266 CACTTTTAATCGGAGCGGAGAAAAATTTTGGTTATGAACATAATTAAGAAGATC 325
Dy	781 TTTTATATTTANATTAATTTTAAATTTTWTATTAATAATTAATAAAAAAAA 840
Oy	326 TCTTGATACATTTTATGAACCTTTTATGGGTTCCGATAGCTCTTGTGCAAG 385
Dy	841 AAAAAAAAAATATATTAATTAATTAATTAATTAATAATAATMAATTTTWTAAAW 900
Oy	386 AGATATCTTGAGAGTCTCTTATCTTGATTT 419
Dy	901 AAAATTTTNAATTTTATTTTATTTATTTATTNMAT 934
RESULT 8	
CNSOLIGB/c	CNSOLIGBP 910 bp DNA linear GSS 01-JUN-2001
LOCUS	Anopheles gambiae GSS sp6 end of clone 06E14 of NotreDame1 library
DEFINITION	from strain PEST of Anopheles gambiae (African malaria mosquito),

ACCESSION	genomic survey sequence.					
VERSION	AL142826					
KEYWORDS	GI:7000944					
SOURCE	GSS.					
ORGANISM	African malaria mosquito.					
	Anopheles gambiae					
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;					
	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;					
REFERENCE	Anophelis.					
AUTHORS	1 (bases 1 to 910)					
TITLE	Genoscope.					
JOURNAL	Direct Submission					
REFERENCE	Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :					
AUTHORS	BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genosco.cns.fr					
TITLE	Web : www.genosco.cns.fr)					
JOURNAL	2 (bases 1 to 910)					
COMMENT	Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J. Direct Submission Submitted (16-FEB-2000) BIMI, Institut Pasteur, 25, rue du Dr. Roux, Paris 75015, France This clone is from an A. gambiae BAC library provided by F.H. Collins and sequenced by Genoscope in collaboration with the Laboratory of Biochem. and Biol. Molec. of Insects, Institut Pasteur.					
FEATURES	Location/Qualifiers					
Source	1..910					
	/organism="Anopheles gambiae"					
	/strain="PEST"					
	/db_xref="taxon:7165"					
	/clone="06E14"					
	/clone.lib="Notredame1"					
	/note="end : SP6"					
BASE COUNT	376 a 83 c 91 g 294 t 66 others					
ORIGIN						
Query Match	8.9%; Score 66.6; DB 17; Length 910;					
Best Local Similarity	43.5%; Pred. No. 0.021;					
Matches 173; Conservative 27; Mismatches 197; Indels 1; Gaps 1;						
OY	89	TTTCGATCAAACTTGGAAAAAGTTAAATTAAATAGTCCAAAAGCCTTTTAAATTT	148			
DB	906	TTATATMTATTTTGATATATATTTTAAATATATAATTAATGAATCATATTAATAATVA	847			
OY	149	CATTATATATGCACATATATACCAGTTAAATTAAGAAGCGTTTTATGAATTAACCTTTAT	208			
DB	846	AAMATTAAMWTTATTTTARNAATTTTATAMNANNWTGTMAKTAFTTNMTWAAT	787			
OY	209	TTTTATTTTGGTAATCTTTTGCTTTTCTAGTTTGGTCAAGATNGATCTTAAAGCAC	268			
DB	786	TTTTTAAWAAATTTNNATTTTNTTATTTTATATATATTTTWWTTTAAATAT	727			
OY	269	TTTTATATCTGGGAGCGGGAAGAAATTTTGGTTATGAACTAATAGAAAGATTCCT	328			
DB	726	TTTTTTTATATWTAAGTAAATTTAAATTTTATTAATTAATTAATTAATTAATTTTWT	667			
OY	329	TGTACCATTTTATGAACCTTTTATAGGGTTCGGGATA-GGTTCTTTTGGTCAAGGAG	387			
DB	666	TATATATATTAAMWMAATANTTTTATTTTATTTTATTTTATTTATTTTNTTAAATTA	607			
OY	388	ATATCCTTGAGAGCTCTTATTTCTTGATTTTGATGCGGTGTGTAAGGTTAATACTTA	447			
DB	606	TAAATAAAAAATTTTWTATTTTAAATWMAATATTAATTTTATATTAATTAATTTTAA	547			
OY	448	CAGGAGCTATTTGGACATCAGAGATTTGATTAATAT 485				
DB	546	AAAAATATATATTTTAAAAATTTTATATWTAATAAW 509				
RESULT 9						
CNS00EVL/C	CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999					
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC:					
DEFINITION	BACR29B23 of RPc1-98 library from Drosophila melanogaster (fruit					

ACCESSION	fly), genomic survey sequence.
VERSION	AL069706
KEYWORDS	AL069706.1 GI:4949849
SOURCE	GSS.
ORGANISM	Drosophila melanogaster.
	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephydroidea; Drosophilidae; Drosophila.
	1 (bases 1 to 1101)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Oseegawa and Aaron Mammocari in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .

FEATURES	Location/Qualifiers
source	1. .1101 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACR29B23" /clone_1ib="RPC1-98" /note="end : 17"
BASE COUNT	419 a 91 c 60 g 299 t 232 others
ORIGIN	
Query Match	8.7%: Score 65; DB 17; Length 1101;
Best Local Similarity	32.9%; Pred. No. 0.03;
Matches 161; Conservative	84; Mismatches 244; Indels 0; Gaps 0;
OY 15	AATAAGATATTGTAAAGAATGTAAGTGAATTTGATGGTTAATCAAGAAGCTCT 74
Db 1099	WMTATWTTTTTTTTTTWTATWTATAMATATATTAATTTWTTATATWTAAMWTTATW 1040
OY 75	ATTGGTAAGCGAATTTTCGATPAACAATTTGAAAAAGTTAAATTAATAGTYCCAAG 134
Db 1039	ATATATTAAATTTWTAATAATATAMWTTAATTAATATATATWTAAMWATATATAMATANAAW 980
OY 135	CCTTTTTRAAATTTGCATPATATGCTACCATATATCCAGTTAAATAAAGGGSTTTTATG 194
Db 979	TATTAAMWMAATWMAATWMAATWMAATTTAATATAMATATAWMAAMWWTTATATTTTTTTTT 920
OY 195	AATAAGTTTTATTTATTTATTTGGTAACTTTTGCTTTTCAGTTTGCTCAAGAT 254
Db 919	TTWTTTWTAATWTTTATATATWTTWMAAAMAMWAAMWATATTTTTTTTWTATATTATATAT 860
OY 255	GATTCTAAAAGCATTTTAATCTGGAGCGGAGAAAAATTTTGGCTTATGAACATAAT 314
Db 859	TAMATTTTTTTTWTATTTWTATATATATWTTATWTTATWTTTWTATATTTATATWTTW 800
OY 315	AAGAAGATTCCTGTACACATTTTATTTGAACCTTTTTTAGGGTGGAGTAGGTCT 374
Db 799	TTWTATTTWATATATAMWTTMWTMMWTAMAAMTATATATWMAATATTTATWTTWATTWTTW 740
OY 375	TTTGCCTCAAGAGATATCCTTGGAGGTCCTATATTTGCAATTTGATGCGGTTGGTATA 434
Db 739	TTTTATTTWTTTAAATTAATATATATWTTWMAATWTTTTTTTTTTTTTTTTTWTATTTMAAAT 680
OY 435	GGTTTAATACTTACAGAGACTTATTGGACATCAAGATTTTGAATATATGCTAAAAAAAA 494

Db 490 AAAAAAAAAATGTAATTAATTTTATKAMWMAAATAATTTTWTWMAAA 549
 Oy 65 AAGAGCTCTATGTAAGCAATTTGAGTACATTTGAAATTAATTAATA 124
 Db 550 AATATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 609
 Oy 125 GTTCAAAAGCCTTTTAAATTCATTATATGCTACATTAATTAAGG 184
 Db 610 WTAAATWMAAATATWTAATATATWTAATWTAATWTAATWMAA 669
 Oy 185 GGTTTATGATTAAGTTTATTTATTTTGTATCTTTGCTTTCTAGT 244
 Db 670 AAAAAAATAAATTTTAAATAAATAAATTTTATTTTWTATWTA 729
 Oy 245 TGCACAGATGATCTTAAGACCTTTTAACTGAGCGGAGAAATTTGTTA 304
 Db 730 WATMAAATATWMAAATWTDKNNNNNNNNNNNNNNNNNNNNNNNN 789
 Oy 305 TGAACATAATAGAAGATCTCTGTACATTTTATGAACCTTTTATGAGG 364
 Db 790 AAAAAAAAAAAAAAAAAAADDMMKAKKKKKKKKKKKKKKKKKKK 849
 Oy 365 GATAGTCTTTTGTCTAAGAGATCTGAGGTTCTTATCTGATTT 419
 Db 850 TTTTTCAGATTTTCTTATTTTCTTATTTTCTTATTTTCTTAT 904

RESULT 12
 CNS020K7/c 1092 bp DNA linear GSS 12-MAY-2000
 LOCUS Tetradon nigroviridis genome survey sequence T7 end of clone
 DEFINITION 222L11 of library G from Tetradon nigroviridis, genomic survey
 sequence.

ACCESSION AL175696.1 GI:7813753
 VERSION AL175696
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetradon nigroviridis.
 ORGANISM Tetradon nigroviridis.

REFERENCE 1 (bases 1 to 1092)
 AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 TITLE Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1092)
 AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1092)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at
 http://www.genoscope.cns.fr/tetradon.

FEATURES
 source
 1. 1092
 /organism="Tetradon nigroviridis"
 /db_xref="taxon:99883"
 /clone="222L11"
 /clone_lib="G"
 /note="Genoscope sequence ID : COAG222CTF06LPI-end : T7"
 BASE COUNT 383 a 169 c 165 g 262 t 113 others

ORIGIN
 Query Match 8.6%; Score 64.2; DB 17; Length 1092;
 Best Local Similarity 35.8%; Pred. No. 0.052;
 Matches 133; Conservative 62; Mismatches 177; Indels 0; Gaps 0;

Oy 6 TTTTCAACAATTAAGATATTTGTTAAAGAATTCATTTGTAATTTATGTTAATCA 65
 Db 949 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 890
 Oy 66 AGAGCTCTATGTAAGCAATTTGAGTACATTTGAAATTAATTAATAG 125
 Db 889 TTAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 830
 Oy 126 TTCCAAAGCCTTTTAAATTCATTATATGCTACATTAATTAAGG 185
 Db 829 WTAAATWMAAATATWTAATATATWTAATWTAATWTAATWMAA 770
 Oy 186 GTTTTATGATTAAGTTTATTTATTTTGTATCTTTGCTTTTCTAGT 245
 Db 769 WMAAATWMAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 710
 Oy 246 GCTCAGATGATCTTAAGACCTTTTAACTGAGCGGAGAAATTTGTTAT 305
 Db 709 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 650
 Oy 306 GAACTAATAAGAAGATCTCTGTACATTTTATGAACCTTTTATGAGG 365
 Db 649 AAAAAATTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 590
 Oy 366 ATAGTCTTTT 377
 Db 589 TTGGCGGAGTTT 578

RESULT 13
 CNS01TL7/c 1038 bp DNA linear GSS 12-MAY-2000
 LOCUS Tetradon nigroviridis genome survey sequence T7 end of clone
 DEFINITION 194L16 of library G from Tetradon nigroviridis, genomic survey
 sequence.

ACCESSION AL166660.1 GI:7804398
 VERSION AL166660
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetradon nigroviridis.
 ORGANISM Tetradon nigroviridis.

REFERENCE 1 (bases 1 to 1038)
 AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 TITLE Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1038)
 AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 1038)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at
 http://www.genoscope.cns.fr/tetradon.

FEATURES
 source
 1. 1038
 /organism="Tetradon nigroviridis"
 /db_xref="taxon:99883"
 /clone="194L16"
 /clone_lib="G"
 /note="Genoscope sequence ID : COAG222CTF06LPI-end : T7"
 BASE COUNT 383 a 169 c 165 g 262 t 113 others

```

source                                1. .1038
                                        /organism="Tetrodon nigroviridis"
                                        /db_xref="taxon:99883"
                                        /clone="194P16"
                                        /clone_1id="G"
                                        /note="Genoscope sequence ID : COAG194DH08LP1-end : T7"

BASE COUNT      366 a    152 c    189 g    145 t    186 others
ORIGIN

Query Match      8.4% Score 62.8; DB 17; Length 1038;
Best Local Similarity 38.8%; Pred. No. 0.092;
Matches 99; Conservative 50; Mismatches 106; Indels 0; Gaps 0;

OY   23 TATGTCTAAAGAATTGGAATTCGTAATTTATGGTTAATCAACAGACTCATTTGGTAA 82
     ||| | | : : : : : : : : : : : : : : : : : : : : : :
Db   813 TDTTTTRRRRRRRRARRRRRRRADRRATTTTTRGRKATRKRAATTTTTRATTTTK 754

OY   83 GCGAATTTCCGATCAACAATTTGAAAAAGTTAAATTTAATAGTCCAAAACCCTTTT 142
     : ||| : | | : | | : | | : | | | | | | | | | | | | | |
Db   753 RATTTTTTTTTRRARATTTTAAWAAATTTKAATTTKATATTTWTTRBAATTTWTMDT 694

OY   143 AAATTCATTAATAGTCACCATATACCAGCTTATAAGGGCTTTTATGCAATAAGTT 202
     ||| | | : | | : | | | | | | | | | | | | | | | | |
Db   693 TTTTCTTTTCKMAAAWAAAAAATTTTAAARAAATTTTAAAAATTTTATTATTAAT 634

OY   203 TTTAATTTTATTTTGTAATCTTTTGTCATTTTCTAGTTTGCACAGATGATCTAA 262
     ||| | | | | | | : | | | | | | | | : | | | | | |
Db   633 TTTAAATTTTTTTWTVMITNTTTTTTTTTTTMMAMTTTITTVTRCTATTTVVMVM 574

OY   263 AAGCAGCTTTTAATCT 277
     : : : : : : : : : :
Db   573 MMAMVTMTKTMTT 559

RESULT 14
CONSOL6E2/c      1204 bp      DNA      linear      GSS 26-JUL-1999ss
LOCUS            Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION       BACN15A12 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
VERSION          AL106628
KEYWORDS         AL106628.1 GI:5622852
SOURCE           GSS.
ORGANISM         Drosophila melanogaster.
AUTHORS          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
TITLE            1 (bases 1 to 1204)
JOURNAL          Genoscope.
REFERENCE        Direct Submision
AUTHORS          Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr)
TITLE            - Web : www.genoscope.cns.fr)
COMMENT          Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

FEATURES             Location/Qualifiers
     source           1..1204
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACN15A12"
                     /clone_1id="DrosBAC"
                     /plasmid="pBelobAC11"
                     /note="end : T7"

BASE COUNT      298 a    172 c    106 g    316 t    312 others
ORIGIN

```

[illegible]

RESULT 15	LOCUS	CNS035N7	576 bp	DNA	linear	GSS 15-MAY-2000
DEFINITION	Tetraodon nigroviridis genome survey sequence pUC-Ori end of clone 21A06 of library G from Tetraodon nigroviridis, genomic survey sequence.					
ACCESSION	AL228940					
VERSION	AL228940.1 GI:7887933					
KEYWORDS	GSS: genome survey sequence.					
SOURCE	Tetraodon nigroviridis.					
ORGANISM	Tetraodon nigroviridis					
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei; Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.					
REFERENCE	1 (bases 1 to 576)					
AUTHORS	Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J.					
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 576)					
AUTHORS	Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.					
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis					
JOURNAL	Unpublished					
REFERENCE	3 (bases 1 to 576)					
AUTHORS	Genoscope.					
TITLE	Direct Submision					
JOURNAL	Submitted (12-APR-2000)					
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at					

